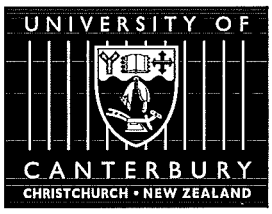
Ring Laser Dynamics

A thesis
submitted in partial fulfilment
of the requirements for the Degree
of
PhD in Physics
in the
University of Canterbury
by
B Tom King



University of Canterbury
1999

To my wife Kirsten

Abstract

The departure of the behaviour of large ring laser gyros from the ideal is examined. A detailed description of the experimental operation of large ring lasers is provided along with several new innovations in equipment layout, data collection and especially in data reduction.

The limits on gyro performance due to noise are investigated. A review of literature regarding the fundamental limit placed on gyro resolution is provided. This limit is due to spontaneous emission in the gain medium of the laser and it is demonstrated that our ring lasers approach this quantum limit. Two entirely independent methods for evaluating the quantum noise induced linewidth are demonstrated to agree well. One of the methods, which uses a second order autoregressive model, is able to make accurate linewidth estimates in sub-second gate times.

A complex model is proposed which accounts for specific observed light scattering phenomena within a ring laser. This model is compared with dual beam data taken from C-I and is able to describe frequency shifts and waveform distortion accurately. The model also performs favourably when describing locking profiles for low rotation rates and externally induced perimeter modulation. When locked to an external signal the ring laser is found to be an extremely sensitive low frequency vibration detector.

The commissioning of a very large (14 m perimeter) prototype ring laser gyro, G0, is described along with a comparison with the smaller (4 m perimeter) gyros C-I and C-II. This prototype has proven to be an invaluable testing ground for designs and techniques to be used on a proposed high precision 16 m perimeter gyro named the Grossring (G).

Acknowledgements

First I would like to thank my supervisor Geoff Stedman for his guidance throughout my time as a graduate student. I would also like to thank Steve Cooper who has always been available for a good talk and who has improved my knowledge of experimental physics immensely.

My family has been extremely patient and supportive while waiting for the completion of this thesis. Thanks Mum, Dad, Dave, Laurence and Gerard.

I am grateful to the Department of Physics and Astronomy for providing excellent technical support, lots of resources and a place to work. I would also like to acknowledge the support of the University of Canterbury Doctoral Scholarship.

I have enjoyed many enlightening discussions while at Canterbury, some of them relevant, most of them not. While I kept up my end of these discussions the other end was usually held up by one or more of Siva Naguleswaran, Andrew Richards, Richard Neutze, Dave Frame, Clive Rowe, Morrie Poulton, Duncan McLeod, John Pritchard, Jason Harris, Doug Wright, Ata Maole, Andy Russell and far too many others to list.

Contents

1. Introduction

This thesis investigates the phenomena that cause the output frequency of a ring laser gyro to differ from that which is calculated from its geometry, orientation and rotation rate. An ideal gyro is dependent only on these three criteria and a real instrument can only approach the ideal if a good understanding of these perturbing phenomena is available.

Stabilising the ring laser output frequency is vital because frequency drift prohibits the detection of small frequency modulations. The ring laser gyros operated at Canterbury have the potential to be extremely sensitive to changes in rotation rate and ultimately may be able to detect fluctuations in the Earth's rotation rate (Ω_E). These fluctuations are of the order of $10^{-9}\,\Omega_E$ which is not far below the lowest stability (5×10^{-6}) achieved on our most stable gyro C-II. Regular upgrades and improvements are continually reducing this stability level. Once the necessary stability is achieved these lasers can replace astronomical observations as the source of Earth rotation information. They will provide rapid turn-around times for rotation measurements and will not be interfered with by weather conditions. The ring lasers are also of interest as test instruments for fundamental physics. A non-reciprocal effect which is applied to the ring laser only has to shift the optical frequency difference between the modes by $\sim 10\mu$Hz in order to be easily detected. This gives a theoretical resolving power of the order of 10^{20}. However, to produce a non-reciprocal effect usually requires a section of the perimeter where the polarisations of the counter-propagating modes are orthogonal. This can only be achieved by introducing optical elements to the cavity which reduces the cavity quality factor and in turn the resolving power of the laser. No such experiments have been performed yet but several experiments have been proposed [21, 32, 109, 110].

Fig. 1.1 provides an illustration of how small variations in the precise location of one of the four mirrors can have a dramatic influence on the output frequency from the gyro. Such sensitivity to operating conditions requires us to achieve a high level of understanding of the ring lasers before their full potential can be realized.

1.1 Background

In 1913 the detection of fringe shifts from rotating interferometers was reported by Georges Sagnac. An historical account of Sagnac interferometers is provided

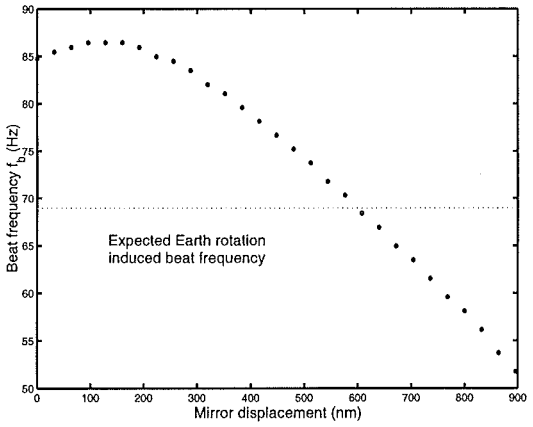


Fig. 1.1: This figure illustrates the ability of one of our ring lasers, C-I, to produce beat frequencies far above and below the expected frequency. The frequency shifts are induced by moving one of the cavity mirrors using a piezo drive. This data was collected on 28th November 1995.

by Anderson et. al. [5]. Modern gyros rely upon the Sagnac effect for their operation and can be found in virtually all applications where once mechanical gyros were used. Optical gyros come in two fundamentally different types, passive and active. Passive gyros are so called because they do not require a gain region to operate. An external laser beam is split with one beam passing around a closed path. The phase shift between the two beams is a measure of the rotation rate of the gyro. This phase shift is scaled by the number of turns so that a physically compact unit may effectively enclose a large area. The greatest limitation of passive gyros results from the difficulty with which the phase shift can be measured. Discussions of this and other passive gyro phenomena are available in several excellent reviews [17, 28, 70].

The work reported in this thesis is concerned only with active ring laser gyros in which two cavity modes are excited, one in either direction; a gain medium sustains lasing of these modes. When in a non-rotating frame the cavity resonant frequencies are equal; rotation lifts this degeneracy. A planar gyro with area vector $\boldsymbol{A}$, perimeter P operating at wavelength λ gives rise to a frequency difference between the two lasing modes

$$f_b = \frac{4\boldsymbol{A}\cdot\boldsymbol{\Omega}}{\lambda P} \tag{1.1}$$

where $\boldsymbol{\Omega}$ is the rotation vector of the gyro. The Canterbury ring laser systems include three gyros. Most of the investigation undertaken in this thesis is concerned with the first ring laser constructed at Canterbury which is named C-I. A second ring laser, C-II, has been constructed, with much higher technical specifications, in a collaboration between the University of Canterbury, Technischen Universität München and Bundesamt für Angewandte Geodäsie. The most recent (January 1998) ring laser is considerably larger with a perimeter of 14 m and was built as a prototype for a proposed 16 m perimeter instrument referred to as the *Grossring* or G. With its status as a prototype we refer to this very large gyro as G0.

All the ring lasers are located at latitude 43°34′37″ south[1] and all operate at the

[1] 1 arc second equates to $\approx$ 31 m at the Earth's surface and all the ring lasers are within this distance of each other.

632.8 nm Neon transition (633.0 nm in vacuum). The theoretically expected beat frequency from, for example, C-I may be calculated as follows. The perimeter is known accurately from the free spectral range frequency, $\nu_{\text{fsr}} = 86.488\,\text{MHz}$, according to $P = c/\nu_{\text{fsr}}$. The period of the Earth's rotation is well known at 0.99726968 days (1 day = 86 400 SI seconds). C-I is slightly rectangular so that the area is not solely determined from the perimeter value but is accurately measured as $A = 0.74996\,\text{m}^2$. These values lead to a beat frequency of 68.724 Hz. A similar calculation for C-II gives a beat frequency of 79.40 Hz. The calculation for G0 is a bit more involved (see §6.6) and gives a value for the beat frequency of 287.75 Hz.

The premise for ring laser gyro operation is that the device gives an oscillating output the frequency of which is linearly related to the rotation rate of the device. In reality this linear relationship is altered (see Fig. 1.1) by the complex nature of the device. This thesis explores these complexities.

1.2 Thesis Layout

A wide variety of subject areas are considered in this thesis as a result of the multifarious nature of the answer to the question that it poses. To this end the thesis contains several chapters that stand alone and may be read in any sequence. These chapters are presented here in Chapters 4 through 6 with Chapters 2 and 3 providing introductory material regarding the operation and analysis of our ring lasers. Chapter 7 concludes the thesis with a summary of the key results along with an appraisal of the achievements made and areas of ongoing research.

The construction and operation of a ring laser is of great importance. In his thesis Ziyuan Li [128] documents much of what is required to arrive at an operating ring laser that unlocks due the Earth's rotation. Further information is contained in many of the local publications [5,21,108,110]. Chapter 2 documents much information that is not formally available. While providing a useful starting point for new members of the group this chapter also serves the purpose of documenting several technical matters that I have developed. Examples of this include the vacuum system and the gas system. The vacuum system described has been rebuilt from an earlier system; the rebuild removed all non-metal fittings (allowing high temperature bake-out), increased the effective pumping speed by a factor of 200 and mounted the system on a trolley allowing it to fulfil a more versatile role. The choice of gas pressure and He-Ne ratio also constitutes a report of my own findings along with examination of contamination rates and the bleed-through method which allows very long operation of the ring laser. Other technical matters such as mirror cleaning, radio frequency plasma excitation and cavity alignment were inherited from the ring operation prior to my joining the group; I have altered these methods slightly for my own purposes.

Chapter 3 is concerned with the variety of numerical methods and techniques

which I have used to extract information from the data recorded from the apparatus. Much of this chapter is concerned with reviewing earlier techniques and explaining the need for faster and more reliable programs in order to extract greater information from the data. The chapter results in the introduction, to the group, of the use of a second order autoregressive model (AR(2)) to describe the ring laser signal and allows the fast and accurate determination of the carrier frequency from the ring laser. This method also proves capable of measuring directly the linewidth imposed on the carrier signal by white noise which originates from the quantum noise intrinsic to the ring laser operation.

The results from the study of the AR(2) technique are used in §4.4.4 where it is shown that two independent methods for determining linewidth agree well. For this reason my investigation of noise in ring lasers is chosen to follow as Chapter 4. The goal of this chapter is to identify the process by which different types of noise limit the performance of ring lasers. Initially many different types of noise are described, all of which may be reduced by appropriate choices of bandwidths, low-noise amplifiers etc. to a level where they are negligible. This leaves quantum noise and environmentally induced frequency drift as the major limits on performance. A review of the current understanding of quantum phase is provided as a prelude to the discussion of quantum noise. In this review it is shown how the conventional quantum limits may be surpassed using squeezed states of light. Also a novel experiment to determine the quantum limits of ring laser performance is proposed in §4.3.5. The conventional approach to quantum noise is presented and it is demonstrated that our lasers operate near the limit imposed by such noise for short measurement times. Further to this, the method of using the Allan variance to characterise gyro performance is discussed. With this method a comparison is made between the local lasers which contrasts the relative merits of the stable design of C-II and the size of G0 with results from C-I. With quantum noise governing performance at short integration times long term stability becomes the key issue. An attempt to combat environmental variations is made which proves successful in improving stability for all integration times where environmental fluctuations prevail; the method gives no improvement where quantum noise is dominant.

The removal of drift due to environmental variations, discussed in Chapter 4, is rudimentary in that a linear model is used to relate pressure and temperature variations with output frequency. Ring lasers are not linear rotation sensors. A complex model is required to describe the interactions between geometry, gain, scattering, mode coupling and so on. Chapter 5 provides a review of existing models where it is seen that several phenomena that are observed in the Canterbury ring lasers are not described by these models. A new model is presented which accounts for those observations made on the local ring lasers. Several tests of the model's validity are made. It is found that the model successfully describes a wide variety of real output profiles. When one of the mirrors is oscillated a locking phenomenon is observed the details of which are qualitatively in agreement with the predictions of the model. The locking behaviour of the ring under different rotation rates is also examined. Although this experiment cannot be performed

locally the predictions are in good agreement with other authors' predictions [45].

Stability of ring laser output may be achieved in several ways. C-II represents an effort to decrease sensitivity to environmental variations through excellent engineering. G0 is a prototype large (14 m) ring laser which attempts to reap the benefits available through scaling laws [108]. While output frequency increases linearly with dimension the backscatter coupling reduces quadratically. Such arguments suggest that considerable improvement may be obtained by scaling up the ring laser. Chapter 6 describes the key technical points of G0. This gyro was successfully commissioned by Dr. Steve Cooper and myself. The raw results from G0 are reported at the end of the chapter. Comparison of the relative Allan deviation profiles of C-I, C-II and G0 has been provided earlier in Chapter 4.

2. Technical Digest

This chapter details the technical aspects of the operation of the group's first ring laser C-I. In order to investigate frequency drift in the ring laser, experimental data is compared with a theoretical model. I describe in some detail the operating conditions under which data is collected so that the appropriate parameters can be included into this model (see Chapter 4). A description of all the equipment, assembly procedures and operation techniques is given starting from the initial placement of the lasing cavity on the baseplate and following through to the condition where the laser is operational and unlocked under the Earth's rotation.

2.1 The Ring Laser Cavity

The chamber in which the laser mirrors are located comprises four housings made locally from stainless steel and four interconnecting quartz or Pyrex tubes. Fig. 2.1 illustrates the essentials of the South-East housing. A variety of different internal diameters, d, are available for the constriction section: d =3.8 mm, 5.8 mm and 7.0 mm. All houses have exit ports which allow access to the eight exit beams of the ring laser. The ports are fitted with anti-reflection coated windows and are sealed with O-rings. Two connections to the vacuum system are made in the southern two housings via ConFlat™ 133 fittings from VARIAN. The four houses are seated on the corners of a 1 m square Zerodur™ [1] baseplate. Grooves have been ground in the Zerodur and cut into the bottom of the housings to locate O-rings which form a vacuum seal between the housing and the baseplate. Another groove is cut in the top of each of the housings which locates a further O-ring and which forms a seal with glass lids. These glass lids allow the mirrors to be viewed from above, so providing a useful diagnostic tool, for example to view the scattered light from particles on the mirrors while the laser is operating. All the O-rings are Viton™ (from the DU PONT CORPORATION) and are sealed with Santovac 5™ diffusion pump oil (made by MONSANTO COMPANY).

[1] Zerodur™ is made by SCHOTT and has a coefficient of thermal expansion which is nominally zero. This coefficient is quoted by the manufacturers for three different grades, 1st, 2nd and 3rd expansion class, as $0 \pm 0.5 \times 10^{-7}$, $0 \pm 1.0 \times 10^{-7}$ and $0 \pm 1.5 \times 10^{-7}$ /C° respectively; these values are only guaranteed in the temperature range 0-50 C°. Our slab is made of 1st expansion class glass.

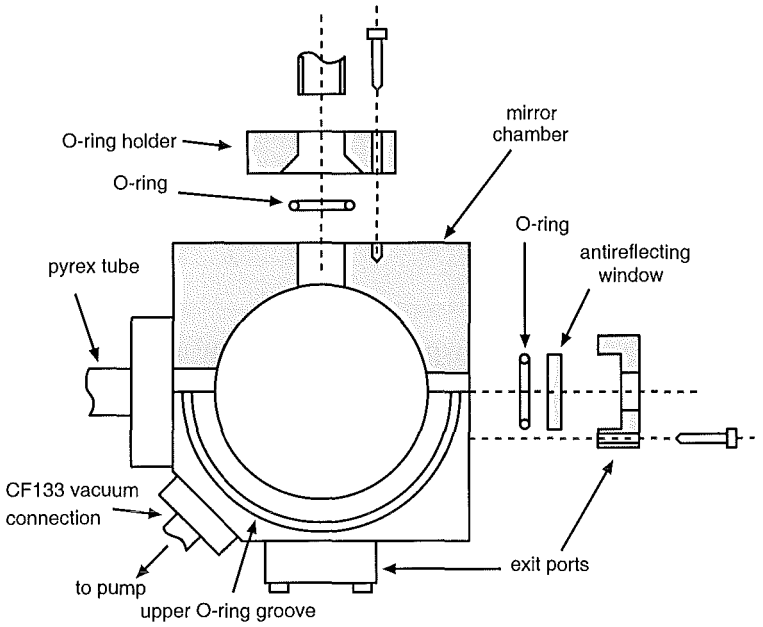


Fig. 2.1: This figure illustrates the construction and assembly of one of the mirror housings. The shaded areas are cut away sections midway up the housing body. With the two connections for the Pyrex tubes, two for the exit ports and the upper and lower seals there are six O-rings for each housing. The drawing is to scale with the box measuring 4 inches along each edge.

2.2 The Vacuum System

In order to lase the cavity must be evacuated and filled with a mixture of helium and neon. A vacuum system, illustrated in Fig. 2.2, has been made and is mounted on a trolley to aid easy removal once the ring laser is in operation. To evacuate the cavity an oil-filled rotary vane pump backs a turbomolecular pump to provide a pumping speed to the vacuum manifold of $50\,l\,s^{-1}$. The two pumps are separated by an oil filter to reduce back-streaming of the backing pump oil into the vacuum manifold. The vacuum system is fitted with several diagnostic tools including a Barocel™ (from EDWARDS HIGH VACUUM INTERNATIONAL), a VacIon™ ion pump from VARIAN (which doubles as a low pressure gauge), a quadrapole mass spectrometer and helium, neon and oxygen gas supplies. The Barocel is a diaphragm capacitance manometer which operates in the range 10^{-3}-13 Torr and which is used to ensure the correct ratio of helium to neon when the laser cavity is filled with gas. The VacIon ion pump serves as a low pressure gauge in the range 10^{-3}-10^{-10} Torr. It is used to determine the quality of the vacuum by measuring the ultimate pressure that is attained. The gauge also acts as a pump. When in normal operation the maximum current is 180 mA. However, if the maximum current is limited to only 17 mA the VacIon ion pump preferentially pumps almost all contaminating gas species over helium and neon thus providing an effective gas scrubbing mechanism. A further measure of the

cavity pressure is provided by the quadrapole mass spectrometer. The spectrometer has the benefit of isolating the partial pressure exerted by species of each given molecular weight. This allows the nature of different contaminating species to be determined and hence provides information about leaks and out-gassing. The helium-neon gas mixture is delivered through high purity gas regulators and then through calibrated leaks. This allows the ratio of the partial pressures of the two gases to be set precisely. Zero grade helium, which has no more than 50 ppm of impurities, is used and the neon is research grade with 1 ppm of impurities.

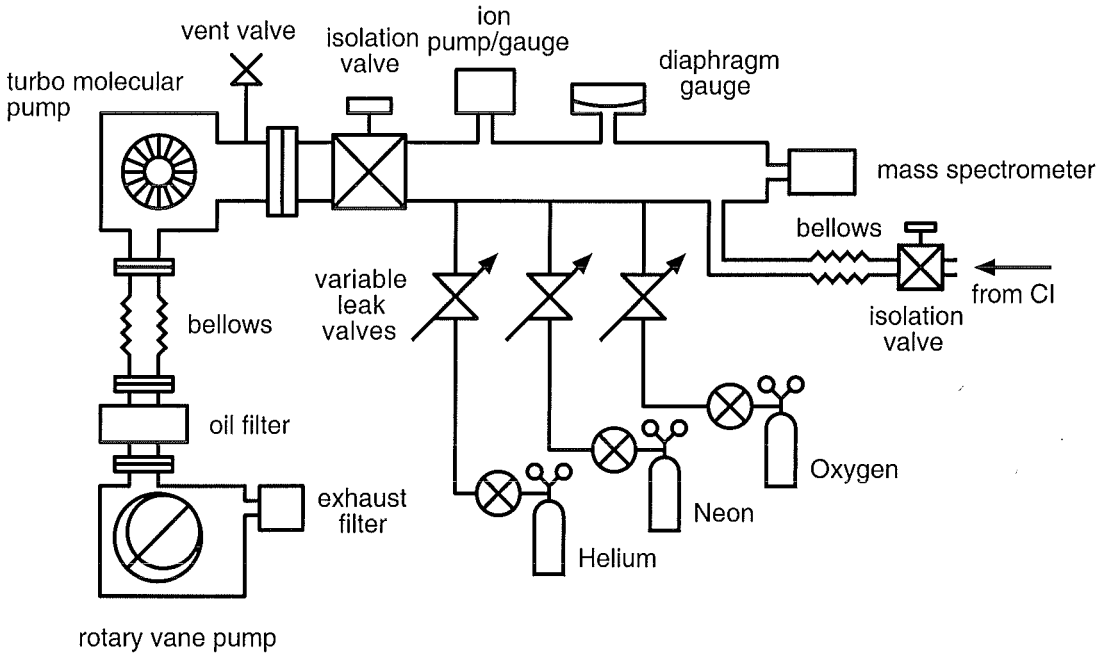


Fig. 2.2: This schematic shows the vacuum system which has been assembled for use on C-I.

The vacuum system contains no O-rings and all components are made from stainless steel. This allows the system to be baked to a temperature of 450°C which, over a period of a few hours, will reduce the partial pressure of hydrocarbon and water contamination to below 10^{-8} Torr. A supply of oxygen is also available which provides a method of cleaning the super-mirrors while *in situ*. This is described by Yamada et al. [127]. This method exposes the super-mirror surface to an oxygen plasma; atomic oxygen reacts with carbon and hydrocarbon deposits on the mirror which are then easily pumped away. These authors report improvement of old mirrors from 1000 ppm total loss to 100 ppm. The treatment is performed in circumstances very similar to those in our ring lasers with the oxygen pressure at 2 Torr and using an rf excitation method. The greatest difference is that our maximum rf power is only 30 W whereas Yamada et al. use 450 W for a period of about 150 minutes. Consequently we must apply the treatment for much longer periods of time. They measure the success of this technique using a ring-down method in conjunction with a tunable dye laser. However, such equipment is not available locally and hence the success of the method can only be

checked at the lasing frequency of the He-Ne laser. For the short-term data sets that are required in Chapter 5 such cleaning methods have not been necessary because the mirrors were cleaned in accordance with the procedures outlined in §2.5 and are hence free of hydrocarbon contamination as is evidenced by the mass spectrometer record.

2.3 Gas Contamination

Long term operation of ring lasers is desirable. If a ring laser is to be run in an *observatory* mode then operation periods of months to years may be required. In such cases the purity of the He-Ne gas mix is of primary interest. There are essentially two strategies that can be adopted. The first is to have a continuous stream of fresh gas flowing through the system and the second is to have a permanent static volume of gas. The first of these options has not been used as standard by the Canterbury Ring Laser Group and introduces the problem of using a large quantity of expensive gas as well as the effects of fluctuations in composition in purity of premixed supplies. In order to maintain a static volume of gas which will support lasing, the rate of contamination by other gases needs to be minimised. The origin of contaminating gases may be from leaks or outgassing from the components of the cavity. Also the He-Ne gas may be absorbed into the volume of the material that defines the manifold of the vacuum cavity.

In a key paper by Martinez [76] many of the phenomena regarding gas contamination are analysed. A description of the decomposition of the fused silica, which forms the walls of the laser tube, is provided. The energy lost by a Ne atom in the lower laser excited state is more than twice that required to break the Si-O bond. This results in a release of oxygen ions into the cavity and leaves a reduced form of silica which tints the glass wall with a brown layer. This tinting of the gain tube has been frequently noted in C-I and has previously been thought to be residual. Further, the release of oxygen into the cavity has been noted while under high vacuum using the mass spectrometer. Unfortunately the rate of release of oxygen while lasing is not easily measured as the mass spectrometer operates only at pressures below 10^{-5} mBar. A subsidiary chamber with an independent high vacuum pump is usually used in such circumstances with a small leak between the chamber to be tested and the subsidiary chamber. Martinez reports that when 250 W of rf power are used the laser continues to operate for 200 hours before oxygen contamination kills the lasing action (nominally about 0.5% oxygen will reduce the laser gain to zero). C-I lases with only a few watts of rf power due to the extremely high reflectance of the cavity mirrors. Therefore the information gained by analysing the gas constituents is not considered worth the outlay of experimenting time. The oxygen that is released into the He-Ne mixture may be gathered by a getter which may be as simple as a container of vacuum baked Aluminium filings. Alternatively the ion pump (which doubles as a pressure gauge) may be left running while the laser operates. This has the effect of pumping most species at about $8\,\mathrm{l\,s^{-1}}$ but pumps the inert helium and

neon at an extremely low rate.

Martinez also noticed that the release of absorbed neon from the silica was sufficient to restart the lasing action. The laser was left running with the usual He-Ne mixture and was subsequently pumped down to 2×10^{-7} Torr for 6-10 hours. Pure helium was injected into the cavity up to 1.5 Torr and the rf field applied. After 30-60 minutes of discharge the cavity started to lase, the necessary neon being desorbed from the wall of the gain tube.

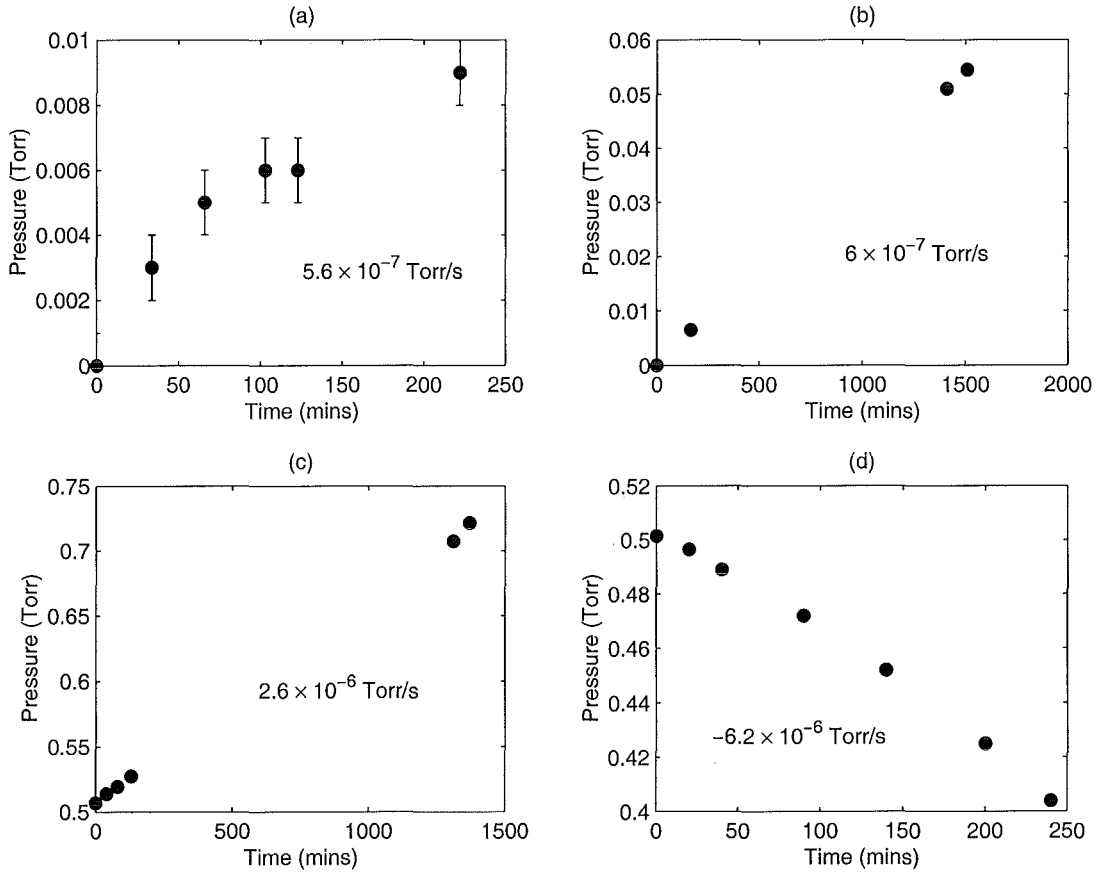


Fig. 2.3: These plots show the outgassing rates measured by the pressure rise technique in which the vacuum pumps are isolated from the ring cavity. In (a) the rf excitation and vacion pump were switched off. The rf excitation was turned on for plot (b). The pressure rise with 0.5 Torr of gas in the cavity is shown in (c) where the vacion was off. When the vacion is subsequently turned on the pressure decrease is recorded in (d). The pressure rise/fall rates are indicated in each case.

Fig. 2.3 shows the results of gas pressure rise measurements performed on C-I. The base line outgassing rate is shown in (a). The possibility that the pressure rise is due to a leak is removed by performing a leak test on the cavity; it is found that there is no detectable intake of the test gas (helium) down to the lowest level detectable by the mass spectrometer. In (b) the slight increase in the pressure rise rate is due to the rf power being applied (10 W forward power) and the mass

spectrometer identifies oxygen as the polluting species. It is assumed that this oxygen is liberated from the glass as suggested in the paper by Martinez [76]. For (c) the cavity was filled with a He-Ne mixture (0.5 Torr) and lasing started. The outgassing rate in the presence of gas appears to be much larger than in its absence. It is assumed that the presence of gas aids the removal of water from the walls of the chamber. When the vacion is turned on and the current drawn limited to 17 mA the pressure in the chamber falls as shown in (d). I have manually countered this pressure rise while collecting data by leaking a premix of He-Ne into the cavity using a fine leak valve. Pressure is successfully kept constant for several days using this method. Further, the effects of gas contamination are reduced by introducing fresh gas to the cavity while selectively pumping the contaminants. The rate at which gas is leaked into the cavity only needs to match the very slow rate at which the vacion pumps helium and neon and hence is not wasteful. This method has only been used manually but could easily be automated using a solenoid driven valve and a simple pressure servo mechanism. There is a difference in the rate of pumping of helium and neon by the vacion. If the above pressure stabilisation method were adopted this would result in a period of settling for the ratio between the two gases; the ratio in the supply volume would be different from that in the laser cavity.

2.4 Mirror Holders

A mirror holder is located in each of the housings. They sit directly on the Zerodur on three small feet as indicated in Fig. 2.4 and have no direct contact with the housings themselves. A cut is made in the stainless steel base; two screws allow forward and backward tilt adjustment by flexure of the base. The super-mirror is held in place by a holding plate which presses the mirror against a wavy washer which acts as a spring. Each mirror is coded with a sequence of markings on the curved surface. This serves to identify the mirror and to allow the repeatable orientation of the mirror in the holder. Appropriate grooves are also milled into the rear of the holder to allow the exit beams through to the exit ports unimpeded. While the tilt adjustments are made manually with the cavity at atmospheric pressure the bulk rotation of the holder about a vertical axis is performed while under vacuum. This is achieved by using an external magnet which provides a torque on the mirror holder via an iron bar attached to the top. While this torque is applied small vibrations in the Zerodur baseplate allow the assembly to rotate. These vibrations are provided by gently tapping the baseplate near the mirror housing; in this manner very fine adjustment of the mirror orientation is achieved. One of the mirror holders has been modified so that a piezoelectric stack may be placed behind the super-mirror. Holes are drilled in the glass lid for electrical lead-throughs and are sealed with Torr Seal™ from VARIAN. This modification affords precise control of the location of the mirror and has been used repeatedly as a diagnostic or as part of normal operation.

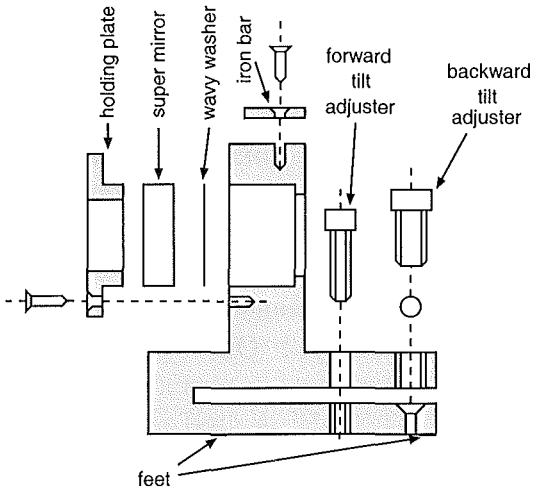


Fig. 2.4: This figure illustrates a schematic of the design of the super-mirror holder. To make the assembly clear the various components are shown in an exploded manner. The super-reflecting surface of the mirror is on the right hand side in this diagram. The base measures $2\frac{1}{2}$ inches in diameter.

2.5 Mirror Cleaning

The supermirrors that form the lasing cavity inevitably become contaminated. There are several sources of possible contaminants including hydrocarbons from the Viton O-rings, impurities in the gas supplies, undesirable out-gassing species etc. After some time (about every 18 months) the mirrors need to be cleaned in order to return them as near as possible to their original state. The cleaning process forms essentially two stages, the first to remove any deposited film on the surface of the mirror and the second to remove particles.

The precise details of the cleaning process cannot be included here as the cleaning agents used and the sequence in which they are applied constitutes commercially sensitive information.

A spinning process is used in conjunction with a wash with the appropriate fluid to dissolve and remove particles from the surface of the mirror. The mirror is mounted on a turntable which may be spun from low speeds up to a maximum of 10 000 rpm. A 6 mW red HeNe laser is directed toward the centre of the mirror, while in the spinning assembly, via a beam expander and illuminates an area of diameter 4 mm. The surface of the mirror is viewed through a microscope and allows continuous feedback regarding the current cleanliness of the middle of the mirror. In §2.6 it is shown that the laser may be readily aligned so that the lasing path remains within this clean area. Transportation of the mirror from the cleaning suite to the ring chamber has to be performed so as to minimise contamination from air borne particles. The mirror is placed in the mirror holder while still in the cleaning suite with a shield covering the reflecting surface of the mirror. This allows the mirror holder assembly to be placed in the mirror chamber without fear of contamination. Only when the glass lid of the mirror housing is ready to be replaced is the shield removed. This cleaning process has proved very successful in restoring the laser mirrors to operating condition.

2.6 Mirror Alignment

Alignment of the cavity is achieved in two stages. The design of the laser requires that the out-of-plane alignment is performed separately from the in-plane alignment. This is facilitated by the fact that out-of-plane misalignments influence only out-of-plane beam walk to first order and similarly for in-plane misalignments. The out-of-plane adjustment screws are contained within the vacuum manifold and can only be adjusted when the system is brought up to atmospheric pressure and the glass lids of the mirror housings removed. Once the out-of-plane adjustments have been made the in-plane adjustments are performed with the cavity under vacuum, as described in §2.4. A GreNe™ is used as a source for

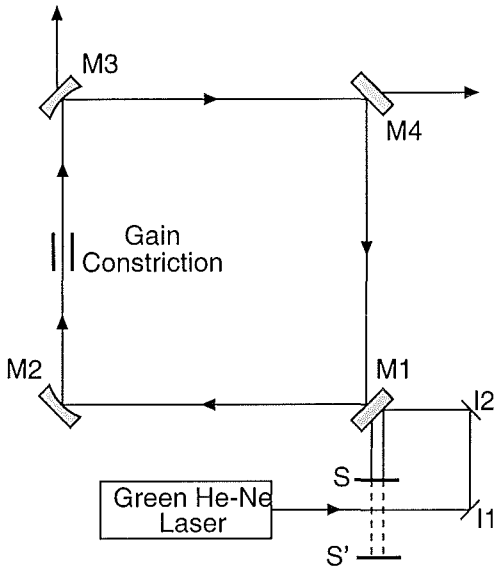


Fig. 2.5: The alignment beam is injected via two mirrors I1 and I2. The beams are visible on the front of the mirrors M1, ... ,M4 through the glass lids and also exiting the cavity through the back of the mirrors. The screen S may be located either close to the exit port or about 2 m away at S'.

the alignment beam[2] because the supermirrors are approximately 50% transmitting at that wavelength. This means that sufficient light is available to perform alignment after two circuits of the cavity. A red HeNe cannot be used as almost no light would pass into the cavity. The arrangement of the components is illustrated in Fig. 2.5. The alignment beam is injected into the cavity and directed approximately to the centre of M1. The beam is located on the mirror by observing the light scattered off the reflecting surface; the alignment spot becomes less easy to locate on the mirror after the mirror is cleaned but can always be located to within ±2 mm of mirror centre. Once the green spot is located on the centre of M1 the transmitted beam's spot is located on M2. Using the injection mirrors I1 and I2 the spot on M2 can be walked towards mirror centre while holding the first spot central on M1. The injection mirrors are not touched again after this and define the axis along which all subsequent alignment is performed.

[2] The MELLES GRIOT green He-Ne laser known as the GreNe™ is used. This lases at 543.5 nm from the $3s_2 \rightarrow 2p_{10}$ Ne transition. These lasers are able to lase on the green transition, which has low gain, without competition from the higher gain 632.8 nm,1.15 μm and 3.39 μm transitions by careful selection of mirror reflectances at these wavelengths.

With the cavity at atmospheric pressure the mirror tilts are adjusted to direct the beam from one mirror to the next starting with M2 and working clockwise to M1. Having completed a full circuit of the cavity the beam (called circuit beam) is aligned with the beam which is initially reflected off the back of the super-reflecting surface (not the back of the mirror substrate) when the alignment beam is initially injected into the cavity (the primary beam). This constitutes a state of coarse alignment and at this point the lids are replaced and the cavity evacuated. Some movement is observed between the primary and circuit beams as the cavity is evacuated but this is small relative to the accuracy to which the spots may be superimposed. A fine adjustment of the in-plane orientations of the mirrors is made before the primary and circuit beams are finally aligned. This is achieved by superimposing the two beams on a screen, S, close to the exit port, and, with adjustment of M1 on a screen, S', 2 m from the exit port. This final step can be performed very accurately with the two spot centres being no more than 1 mm apart.

In summary, the above procedure results in alignment beams that are within 2 mm of the centre of the mirrors and primary and circuit beams that are aligned to within 1 mrad. It was not initially clear that this degree of alignment would provide a stable lasing path located near the mirror centres. It is desirable that the lasing path is close to the mirror centre. The mirror cleaning technique that is used can only guarantee that a circle of diameter 4 mm in the centre of the mirror is properly cleaned (see §2.5). Further, the manufacturer only guarantees the extremely low total losses for a small area ($\sim$ mm^2) at the centre of the mirror. In order to determine whether the degree of alignment is sufficient a beam tracking program has been written (by Dr. Steven Cooper) which allows the location of alignment beams on mirrors and external screens to be compared with the location of the lasing path. Two separate numerical investigations have been performed. The first investigation does not consider the alignment beam but determines how much a single mirror can be rotated, with all the others perfectly aligned, before a lasing spot moves intolerably close to the edge of one of the mirrors. With the reflecting surface taken as 10 mm in radius it is found that all the misalignment angles that kill the lasing action are about 3-4 mrad. This is encouraging as this is greater than the alignment accuracy that can be achieved by observing the green spots on the mirrors. However, this investigation is incomplete as it does not take into account the effects of all the mirrors being partially misaligned and the high degree of accuracy with which the primary and circuit beams may be aligned externally. The second investigation follows the above alignment procedure numerically. The initial orientations of the mirrors and alignment beam are randomised and the alignment procedure followed. When an adjustment locates a spot on the appropriate mirror to within 2 mm of the mirror centre the adjustment is stopped and the next step started. The two exit beams from M1 are also aligned to within the 1 mrad accuracy allowed and thus the entire alignment procedure is reproduced numerically. Finally, the locations of the lasing spots are found and none of them are displaced from mirror centre by more than 2.5 mm. The alignment procedure has been performed successfully

many times on C-I and the locations of the lasing spots (which can only be determined by eye) in all cases are consistent with the above result. Thus it is concluded that the procedure is entirely sufficient for the ring laser's requirements.

2.7 Gas Mixture

A He-Ne laser will lase for many different ratios of helium to neon and at many different pressures. It would seem then that the choice of operating ratio and pressure is somewhat arbitrary. However, in order to make the laser as insensitive as possible to nonrotational environmental effects I have performed output power versus operating pressure measurements illustrated in Fig. 2.6. The measurements are all taken with the laser in exactly the same geometry and with the rf power kept constant throughout at 10 W nett forward power. The cavity is initially filled with a known ratio of helium to natural neon up to a pressure of 2 Torr. The discharge is then initiated and gas gradually bled out of the system using the bellows valve at the top of the turbo pump. Once lasing is observed a series of measurements of combined beam output powers is made at decreasing pressures until lasing ceases. The power jumps that are seen in some of the plots are almost certainly the result of bistability in the plasma which are frequently observed. These bistabilities only appear at high power and do not seem to be important when operating the ring laser at the very low powers associated with single longitudinal mode. In order to make the system less sensitive to pressure drift (atmospheric or outgassing) a stationary point on one of these curves is desirable. Typically I have chosen a ratio of 10:1 and a pressure of 0.7 Torr. The location of the stationary point decreases with increasing proportion of neon. In choosing a ratio of 10:1 I have maximised the pressure which reduces the effects of gas contamination. These characteristics are dependent on the internal diameter of the gain tube (in this case 7 mm) and will have to be repeated for different gain tubes.

2.8 Radio Frequency Plasma Excitation

The plasma excitation scheme and the output power servo are discussed in tandem. The arrangement is shown in the schematic provided in Fig. 2.7. The rf signal generator provides a small GPS stabilised signal to the rf amplifier. This signal is usually somewhere between 50 and 70 MHz and may be varied to optimise the operating conditions. Optimal conditions are taken to be those where the least amount of forward power is required to achieve lasing. This condition is chosen so that the servo has the maximum range to operate in; the gradual contamination of the lasing gas by out-gassing requires an increase in the applied rf field in order to give a constant output irradiance. When the maximum power of 30 W is reached and lasing no longer occurs then the laser requires a refill of

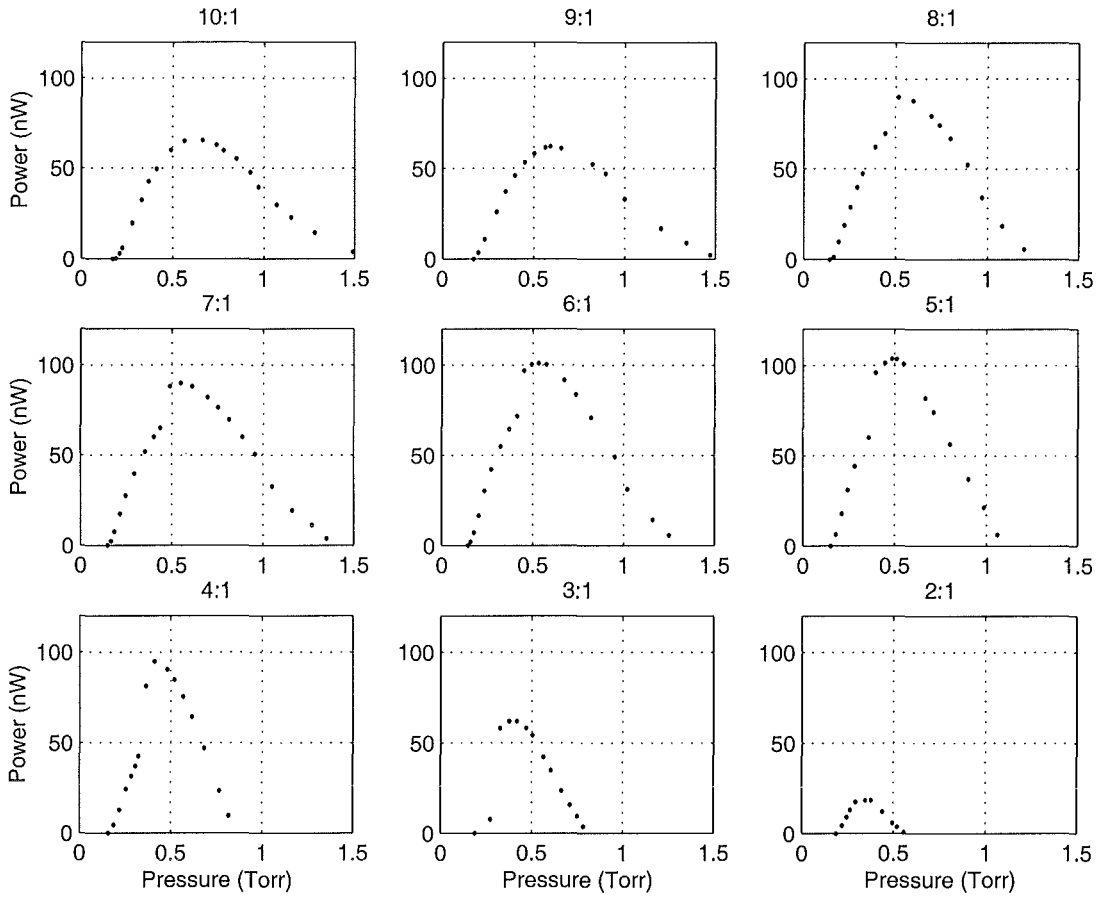


Fig. 2.6: Power versus pressure curves are shown here for a variety of different ratios of helium to natural neon. I typically operated the laser at a ratio of 10:1 with 10 W of nett forward rf power and at a pressure of 0.7 Torr where the output power is independent of pressure. The discrete jump in power at about 0.3-0.5 Torr in several of the profiles is attributed to a bistability in the rf discharge.

gas. The output from the amplifier is coupled to the inductive excitation unit via a matching circuit illustrated in Fig. 2.8. This circuit allows the output from the amplifier to be coupled efficiently to the plasma. The combined output beams are monitored using a PMT with the output current being passed to a spectrum analyser and a transimpedance amplifier. The spectrum analyser allows inspection of the free spectral range of the laser cavity which is located at 86.5 MHz. The absence of any signal at this frequency while lasing indicates the presence of only a single longitudinal mode (see §2.9 for transverse mode analysis). This is the desired operating condition because the model of Chapter 5 describes coupling only between two counter-propagating modes. Although the servo maintains the output power at a constant level it remains possible for thermal drift to alter the operating parameters so that a second longitudinal mode oscillates. This may be understood by considering the initial condition where the laser is operating very close to the middle of the atomic transition line and in single longitudinal mode. A change in the perimeter of the lasing path (possibly thermally induced)

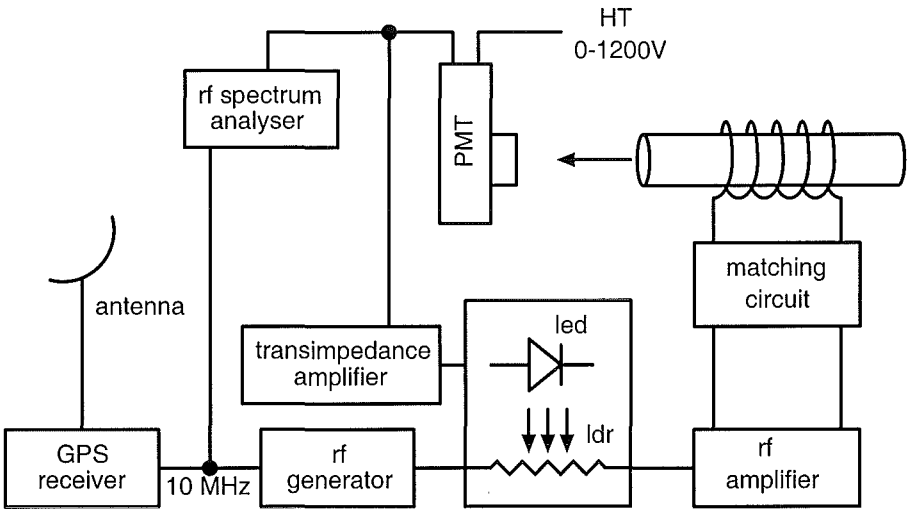


Fig. 2.7: This schematic shows the intimate connection between the rf supply to the plasma and the servo control mechanism.

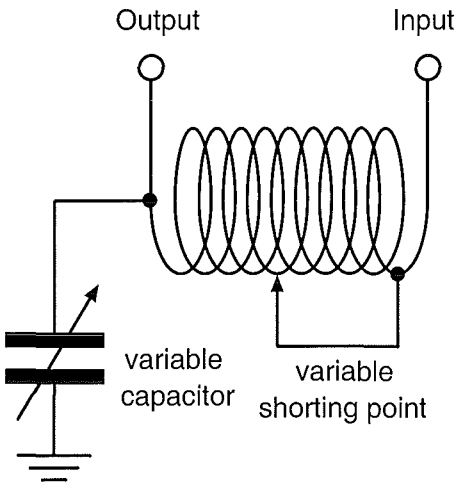


Fig. 2.8: This figure shows the simple rf matching circuit which comprises a single variable capacitor and an inductive coil which may be shorted at several discrete locations along its length.

would cause the lasing frequency to move from line centre. Hence, neglecting the effects of mode competition for now, that mode would receive less gain per pass. The servo mechanism would respond to this decrease in the output irradiance by increasing the rf power delivered to the plasma and hence raising the entire gain profile seen by the cavity. The possibility exists that a second mode may then be allowed to lase despite the output power being held constant. For this reason the spectrum around 86.4 MHz is monitored during data collection to be sure that only a single mode oscillates throughout. No automatic system has been available for this task and hence it was required that the spectrum analyser be observed manually during data collection. This is acceptable for the short data sets (of order 15 mins) which are analysed in Chapter 5 but for long term data acquisition the process will need to be automated.

The current from the PMT is passed to a transimpedance amplifier which provides

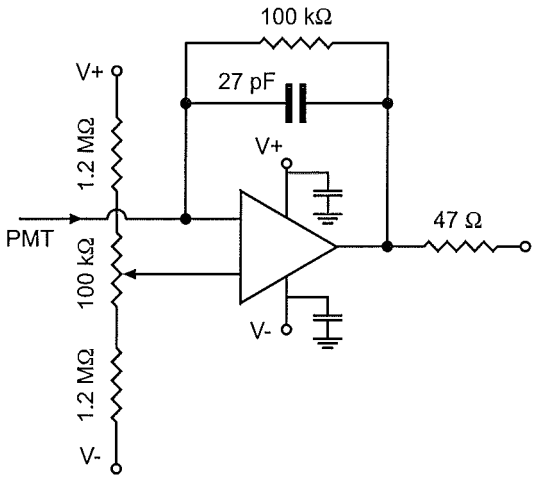


Fig. 2.9: The transimpedance amplifier with R-C output characteristic; the output passband is at about 370 kHz. In order to ensure that the absence of any input light to the PMT corresponds to zero volts from the amplifier, a zero-adjust is provided in the form of a set voltage to the positive input of the op-amp.

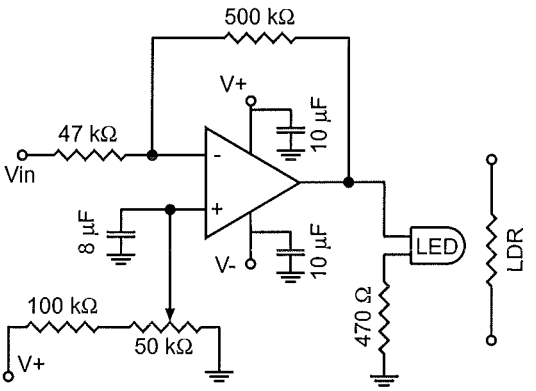


Fig. 2.10: Maintaining a constant output irradiance from the ring laser is achieved using this servo circuit.

a voltage proportional to the output irradiance of the combined beams and which is subsequently fed to the power servo-circuit (see Fig. 2.9). The servo contains a light emitting diode (LED) the brightness of which is governed by the input voltage.

One PMT is dedicated to the stabilisation of the output irradiance and is usually set to observe interferometrically combined beams. A separate PMT observes the output from a different corner and has been used to look at single beam or combined beam outputs. This PMT also converts the output irradiance into a voltage via a transimpedance amplifier and then passes the output to an antialias filter prior to analogue to digital conversion in the data acquisition unit. This data stream constitutes the final ring laser output. Only the simplest detection scheme has been described in this section. A description of a more complex, calibrated dual channel detection scheme is described as part of the ring laser modelling process in chapter 5.

2.9 Transverse Mode Structure

In principle $TEM_{0,0}$ is the cavity mode that most efficiently utilizes the gain provided by the discharge. Operation of the gyro requires only a single longitudinal mode to oscillate. This is achieved by reducing gain until only a single longitudinal mode is provided with sufficient gain to overcome cavity losses. Under these conditions only a single transverse mode can oscillate and this is expected to be $TEM_{0,0}$. However, C-I is often observed to lase in transverse modes other than $TEM_{0,0}$ as indeed has C-II. Several possibilities may give rise to this phenomenon. Diffraction from the body of the laser may reduce the efficiency of one mode over another especially at the gain constriction which is the narrowest part of the cavity. Also imperfections and scatterers on mirrors may provide a loss mechanism for one transverse mode while another may have a node near that point and there would be reduced loss; the second mode may be preferred over the first. Along the same theme the gain profile in the discharge is known to dip in the centre under certain conditions. This would favour modes with lobes off axis, that is, not $TEM_{0,0}$. These phenomena are being investigated by D P McLeod [77].

For the purposes of the model described in this chapter there is no need for a particular transverse mode to lase; the only requirement is that the cavity oscillate in a single longitudinal mode. In Fig. 2.11 a progression of CCD images is shown of the beam spot which are taken as the discharge arrangement is moved in even steps along a 5 cm section of the gain tube. This illustrates that with all cavity parameters kept constant and no change in the supply power different transverse modes oscillate depending on the location of the discharge. Similar mode pregressions are observed as the supply power is varied. In order to ensure that the same transverse mode oscillates throughout operation the CCD camera is used to observe the beam spot profile continuously with a frame rate of about 2 per second. If the transverse mode does change during data collection the calibration procedure (see §5.2.2) is invalidated.

If circumstances arise that demand a particular transverse mode to oscillate then this may easily be accommodated on C-I. Location of the discharge arrangement and changes in the precise cavity orientation have been found to provide any transeverse mode $TEM_{i,j}$ where $i \leq 6$ and $j \leq 3$. An illustration of $TEM_{0,1}$ oscillation is provided in Fig. 2.12. The changes in the cavity orientation are achieved by rudimentary flexing of the cavity arms but proves to be effective. In Chapter 5 data collection and examination is described which does not have any particular demand on which transeverse mode oscillates. With this freedom I have chosen $TEM_{5,1}$ which has the desirable feature of passing cleanly through the PMT photocathode window without grazing any of the internal workings.

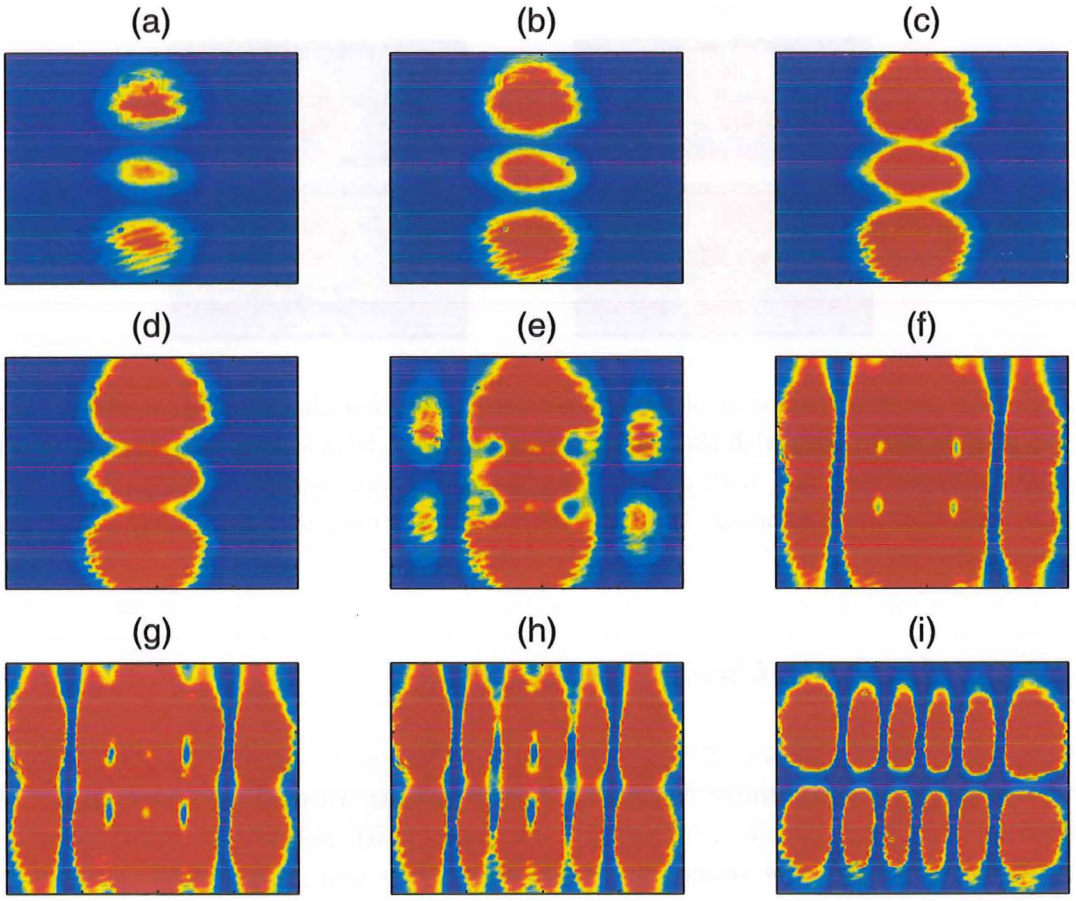


Fig. 2.11: A progression of images of the lasing spot taken using a CCD camera with a 2.56 mm square chip as the discharge position is moved. The images show the progression from $TEM_{0,2}$ to $TEM_{5,1}$. Note that in (e),(f),(g) and (h) both transverse modes oscillate simultaneously.

2.9.1 Transverse Mode Analysis

A considerable part of my investigation into the performance of C-I has required the gathering of peripheral information about the laser. So far the process by which the laser is made operational has been described. Here I consider the radio frequency spectrum observed in the PMT current that is acquired once lasing has been achieved. As the analysis here shows, several important pieces of information are provided from these observations. Firstly the transverse modes which are lasing may be determined much more sensitively than when using a CCD camera. Secondly an accurate measure of the curvature of the super-mirrors is possible; this is important for example when mode matching with an external cavity.

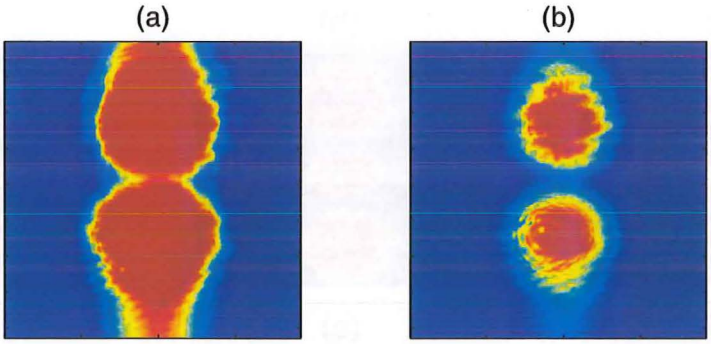


Fig. 2.12: A demonstration of mode selection by varying the cavity geometry. These two images were taken with the supply power high, (a), with several longitudinal modes oscillating and low, (b), with a single longitudinal modes oscillating. Both images are taken with the same exposure time but for (a) an OD2 neutral density filter attenuated the beam.

2.9.2 Analytic Cavity Resonances

In essence the frequency difference between different transverse modes is determined simply by summing the phase due to propagation around the cavity, including the Guoy phase, and equates the total with an integer multiple of 2π in order to satisfy the boundary conditions. The transverse mode analysis of

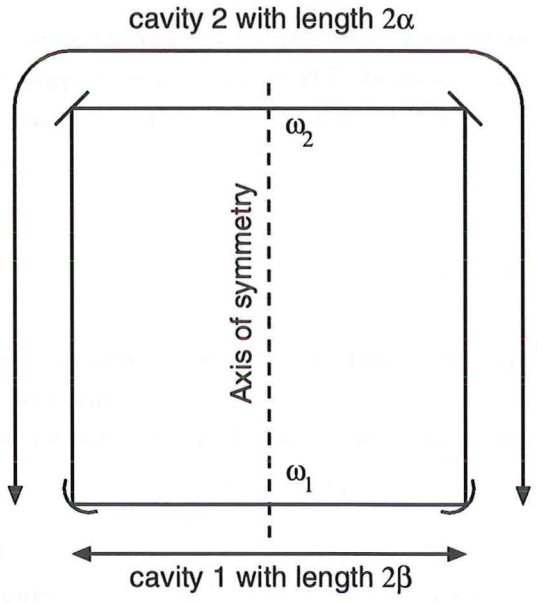


Fig. 2.13: This diagram shows how the cavity is divided into two sections which each have their own beam characteristics.

the ring cavity depends on the placement of flat and curved mirrors. For the measurements taken here the cavity was in a $\begin{smallmatrix} F & F \\ C & C \end{smallmatrix}$ configuration (F denotes a flat mirror and C a curved mirror and the labelling scheme identifies the mirrors in sequence so that $\begin{smallmatrix} F & C \\ C & F \end{smallmatrix} = \begin{smallmatrix} C & F \\ F & C \end{smallmatrix}$). It is convenient to split the cavity into two sections

as indicated in Fig. 2.13. The on-axis ($x = y = 0$) phase evolution of a transverse mode $u_{n,m}(z)$ with a given beam waist is [104, p.686]

$$u_{n,m}(z) = \left(\frac{2}{\pi}\right)^{\frac{1}{2}} \frac{H_n(0)H_m(0)}{2^n n!\omega(z)} \exp\left(i(n+\tfrac{1}{2})\Psi_x(z) + i(m+\tfrac{1}{2})\Psi_y(z) - ikz\right) \tag{2.1}$$

where z is measured from the waist, with $\omega(z=0) = \omega_0$, and H_j is the j^{th} Hermite polynomial. The phase term which relates to higher order transverse modes is $\Psi(z) = \tan^{-1}(z\lambda/\pi\omega_0^2)$ and it is this term that must be treated differently for each partition of the cavity and also for the in-plane (x) and out-of-plane (y) components because of the astigmatism experienced by each of these components. The waist parameters in the two sections of the cavity are found using the complex radius of curvature method. This curvature is defined at the waist as $p_0 = \pi i\omega_0^2/\lambda$ where propagation is performed according to $p(z) = p_0 + z$ and focusing at a mirror of focal length f is performed according to

$$\frac{1}{p_2} = \frac{1}{q_1} - \frac{1}{f} \tag{2.2}$$

with the pellicle limit assumed throughout. After being propagated around the cavity the resulting complex curvature is equated with the initial curvature in order to satisfy the boundary conditions. The manipulations are performed with MATHEMATICA® and give the results:

$$\begin{aligned} \frac{\pi\omega_{1x,1y}^2}{\lambda} &= \left(\frac{(f_{x,y}-\beta)(-\alpha\beta + f_{x,y}(\alpha+\beta))}{f_{x,y}-\alpha}\right)^{\frac{1}{2}} \\ \frac{\pi\omega_{2x,2y}^2}{\lambda} &= \left(\frac{(f_{x,y}-\alpha)(-\alpha\beta + f_{x,y}(\alpha+\beta))}{f_{x,y}-\beta}\right)^{\frac{1}{2}} \end{aligned} \tag{2.3}$$

with $f_x = f\cos\theta$, $f_y = f/\cos\theta$ and $\theta = \pi/4$. The subscripts refer to the cavity section and the in-plane and out-of-plane axes. This leads to the expressions for the phase terms

$$\begin{aligned} \Psi_{1x} &= 2\tan^{-1}\left(\frac{\beta\lambda}{\pi\omega_{1x}^2}\right) & \Psi_{2x} &= 2\tan^{-1}\left(\frac{\alpha\lambda}{\pi\omega_{2x}^2}\right) \\ \Psi_{1y} &= 2\tan^{-1}\left(\frac{\beta\lambda}{\pi\omega_{1y}^2}\right) & \Psi_{2y} &= 2\tan^{-1}\left(\frac{\alpha\lambda}{\pi\omega_{2y}^2}\right) . \end{aligned} \tag{2.4}$$

In order to complete an integral number of cycles in a full orbit of the cavity the wavelength must satisfy the equation

$$(n+\tfrac{1}{2})(\Psi_{1x}+\Psi_{2x}) + (m+\tfrac{1}{2})(\Psi_{1y}+\Psi_{2y}) - \frac{2\pi P}{\lambda} = 2q\pi \tag{2.5}$$

where q is an integer. The transverse mode is often referred to as $\text{TEM}_{n,m,q}$ [104, p.926] where n and m specify the transverse mode and q, usually a very large

integer, specifies the longitudinal mode. The beat frequency between the two transverse modes $\mathrm{TEM}_{n_1,m_1,q_1}$ and $\mathrm{TEM}_{n_2,m_2,q_2}$ may be deduced with a little manipulation to be

$$\Delta f = \left(\frac{\Psi_{1x} + \Psi_{2x}}{2\pi} \Delta n + \frac{\Psi_{1y} + \Psi_{2y}}{2\pi} \Delta m - \Delta q \right) \nu_{\mathrm{fsr}} \tag{2.6}$$

where $\nu_{\mathrm{fsr}} = c/P$ is the free spectral range frequency of the cavity, $\Delta n = n_1 - n_2$ and $\Delta m = m_1 - m_2$. Eq. (2.6) shows that many pairs of transverse modes will have degenerate beat frequencies, for example the pairs of modes $\mathrm{TEM}_{2,3,q}$, $\mathrm{TEM}_{1,2,q}$ and $\mathrm{TEM}_{1,1,q}$, $\mathrm{TEM}_{0,0,q}$ have the same beat frequency.

2.9.3 Measured Cavity Resonances

While lasing, the beam is observed directly using a CCD camera and indirectly using an rf spectrum analyser which takes the PMT current as input. Initially a knife edge was used to ensure that the orthogonal cavity modes would mix and later it was found that imperfections in the detection system were sufficient to provide mode mixing. The spectrum analyser was set to a bandwidth of 1 kHz and the line centre of frequencies attributable to the laser were recorded. Several spurious lines were rejected as they did not disappear when the beam was occulted. The resulting frequencies shown in Table 2.1 are projected down to a

Frequency (MHz)	Decomposition
17.810	
29.709	
32.802	
53.688	ν_{fsr} - 32.802
68.678	ν_{fsr} - 17.810
86.488	ν_{fsr}
104.297	ν_{fsr} + 17.810
116.210	ν_{fsr} + 29.709
119.296	ν_{fsr} + 32.802
140.168	$2\nu_{\mathrm{fsr}}$ - 32.802
155.171	$2\nu_{\mathrm{fsr}}$ - 17.810
172.976	$2\nu_{\mathrm{fsr}}$
193.523	$2\nu_{\mathrm{fsr}}$ + 20.553
241.659	$3\nu_{\mathrm{fsr}}$ - 17.810

Table 2.1: Observed frequencies which are attributable to the laser cavity. The projection of the observed frequency into the interval 0 to $\nu_{\mathrm{fsr}}/2$ is indicated. While these measurements were taken the two transverse modes $\mathrm{TEM}_{5,1}$ and $\mathrm{TEM}_{0,2}$ were clearly visible on the CCD camera.

frequency window between 0 and $\nu_{\mathrm{fsr}}/2$ by subtracting integer multiples of the free

spectral range and taking the modulus. This manipulation is possible because transverse modes with the same m and n index may be separated by multiples of $\nu_{\rm fsr}$ which correspond to different q indices. The number of frequencies under consideration reduces to just four: 17.810 MHz, 20.553 MHz, 29.709 MHz and 32.802 MHz. In this way a spectrum is formed which may be used to identify which modes are lasing at a particular point in time. A second set of results is reported in Table 2.2 where $\mathrm{TEM}_{4,0}$ and $\mathrm{TEM}_{0,1}$ are clearly seen to be lasing using the CCD camera. Matching the observed frequencies with the frequencies

Frequency (MHz)	Decomposition
3.078	
17.805	
20.882	
65.601	$\nu_{\rm fsr}$ - 20.882
68.684	$\nu_{\rm fsr}$ - 17.805
83.411	$\nu_{\rm fsr}$ - 3.078
86.488	$\nu_{\rm fsr}$

Table 2.2: Results obtained when $\mathrm{TEM}_{4,0}$ and $\mathrm{TEM}_{0,1}$ are seen to be present on the CCD camera. The spectrum for these results is 3.078 MHz, 17.805 MHz and 20.882 MHz.

predicted by Eq. (2.6) is performed in a few steps. Initially the frequencies expected for the modes that are known to be present are evaluated using the value of 3 m for the super-mirror focal length which is provided by the manufacturer. It is found that $\nu_{5,1,q+1} - \nu_{0,2,q}$ is 30.071 MHz and $\nu_{4,0,q+1} - \nu_{0,1,q}$ is 3.122 MHz both of which equate well with frequencies observed in the corresponding spectra, namely 29.709 MHz and 3.078 MHz. The assumption is made that these are indeed the correct frequencies and the value for the focal length is varied to optimise the fit between the predicted and measured frequencies. When the fit is optimal the value for the focal length of the curved mirrors is 3.011 m. Fig. 2.14 illustrates the predicted frequencies for many values of Δn and Δm alongside the spectra for the two different data sets. The predictions compare well with the measured spectra and allow the determination of which transverse modes are lasing. The uncertainty in the estimate for the focal length is provided by varying the length β where $4\beta + 4\alpha$ is the perimeter, known to high accuracy from the value for the free spectral range. The uncertainty in β (0.449 m) can be no more than 1 mm and when varied by this amount the focal length at optimum fit varies by the same amount so that $f = 3.011 \pm 0.001$ m. This uncertainty estimate is only rudimentary as the path that the resonant mode follows will in general be trapezoidal and not necessarily planar.

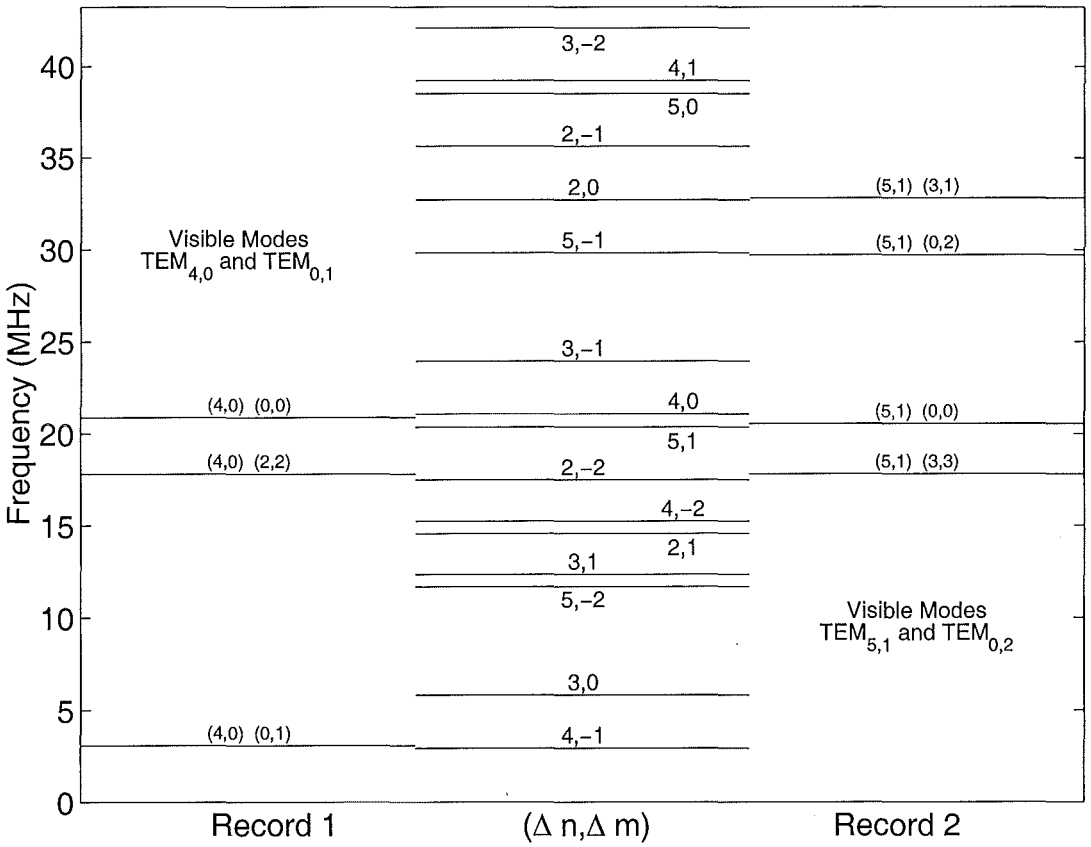


Fig. 2.14: A comparison between the spectra observed on two separate occasions and the theoretical spectrum.

2.9.4 Discussion

The method described here provides a sensitive technique for determining which transverse modes are lasing. This is important for ensuring that only a single mode oscillates. It is possible to have two transverse modes oscillating but with no signal seen at the free spectral range. At present the absence of a signal at this frequency is regarded as sufficient to assure single mode operation. In addition this analysis provides an accurate estimate of the curvature of the super-mirrors in the cavity although the calculations will need to be reworked a little for cavity configurations other than $\begin{smallmatrix} C & C \\ F & F \end{smallmatrix}$.

3. Methods of Analysis Used by the Ring Laser Group

Much of the data analysis performed on ring laser data has been dependent on the Fourier transform and associated filtering and windowing techniques. This chapter explores some of the many alternative methods of analysis that have been used to retrieve as much information as possible from the streams of data taken from the Canterbury ring lasers. A review of the methods used by the Ring Laser Group is also provided. The primary contribution made here is the introduction of the autoregressive technique which provides millihertz frequency resolution from data segments that are less than a second in duration. This super-resolution is only possible as a result of using *a priori* information. This technique also gives an accurate estimate of linewidth which is of particular importance to the noise analysis provided in Chapter 4. The accurate determination of the output frequency on short time scales is an important part of the thesis because frequency fluctuations over short periods of time give us a great deal of information about the ring laser performance. This is exemplified by the Allan variance discussed in Chapter 4. This shows interesting behaviour of the output frequency on time scales from 1 s to 300 s. Had Fourier techniques been used then millihertz resolution would only be possible for data sets of duration more than 1000 s and hence information about the power laws apparent in the Allan variance plot (see Fig. 4.12) would be missed.

3.1 Introduction

A full characterisation of the ring laser behaviour requires the collection of all possible information regarding operating conditions in order to understand better the behaviour of the device. The primary data required for our purpose is an evenly spaced series of samples of some combination of the irradiances of the separate output beams and the interferometrically combined beams. Further information is gained from the logging of the radio frequency spectrum around the free spectral range of the cavity and (from a simultaneous record) of the environmental conditions under which the ring laser is operated. Many diverse records are thus taken. An appropriately diverse range of techniques is required to glean as much information from the collected data as possible. Fig. 3.1 illustrates the diversity

of techniques that are available for the purpose of extracting information from ring laser data. When selecting a data processing technique decisions have to be made regarding which parameters are to be estimated. In Fig. 3.1 the decision path emanates from the data stream in the centre. The flow chart includes a

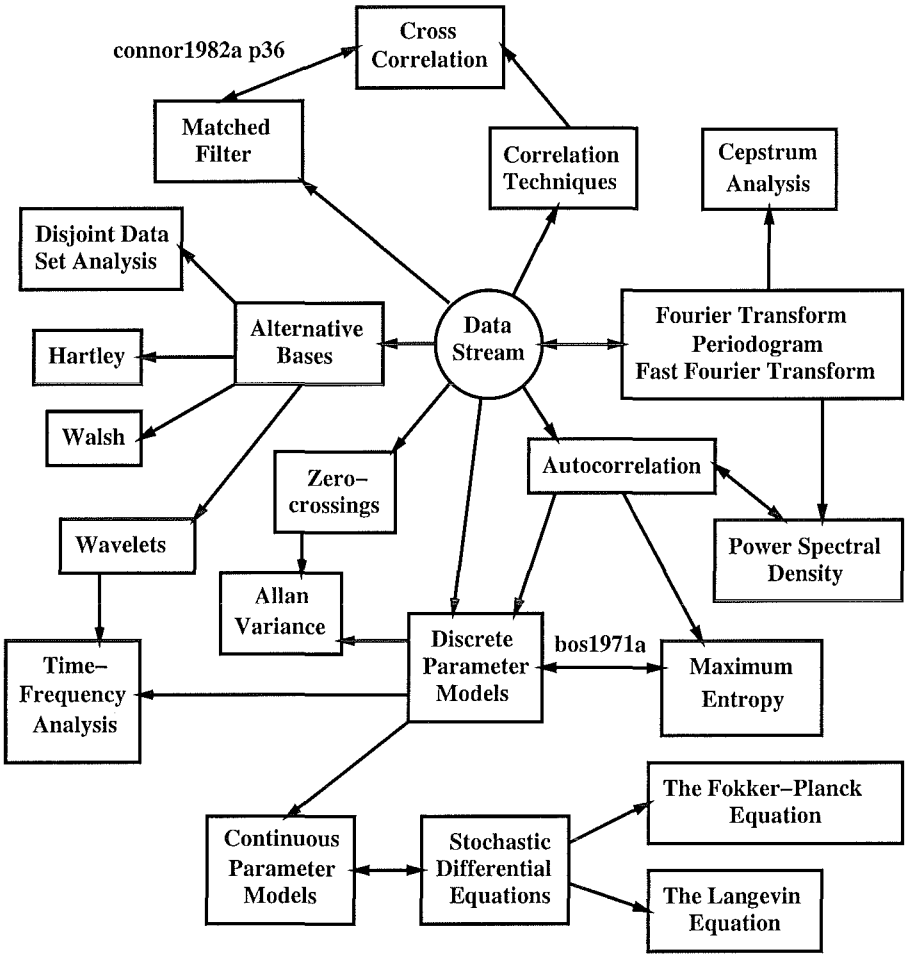


Fig. 3.1: A flow chart that shows how the various methods of analysis most relevant to the ring laser data stream are interrelated. Although this flow chart could be extended indefinitely, it has been truncated at or about the point where anything new would not give significant benefit to ring laser data analysis.

few techniques which are not directly discussed here but whose location in the scheme is useful to know. A double headed arrow on the chart indicates that a particular step is reversible; for example the step of taking the Fourier transform of a data set is reversible but once the modulus is taken the phase information is lost and the original data cannot be recovered.

This chapter begins by documenting the methods that have been used prior to my joining the group and then gives an account of the additions that I have made. Almost all analysis in the past has been based on Fourier techniques, the essential details of which are provided in §3.2, which also thus gives a context for the subse-

quent work. This section highlights those techniques actually used by the group. The main thrust of this chapter is to demonstrate that Fourier techniques do not provide all the information that is contained within the data. Often alternative analyses to Fourier techniques may provide more accurate estimates of desired parameters and occasionally a different technique may provide a completely new perspective not seen at all using the Fourier methods. As an illustration of this an alternative integral transform method called the Hartley transform, quite similar to the Fourier method, is included in Appendix A. This transform provides the same spectral information as the Fourier spectral estimate but presents the information differently and hence provides an alternative perspective.

The analytic signal method has also been used in conjunction with the Fourier analysis. This method uses the Hilbert transform to estimate the instantaneous phase and hence express the real data signal as a complex signal. §3.2.4 explores this method and introduces the idea of using a digital filter to estimate the instantaneous phase, a method which is computationally more efficient; the Hilbert transform relies on two fast Fourier transforms and hence is at best an O(Nln(N)) process whereas the digital filter is always an O(N) process. An estimate of signal phase as a function of time provides access to the modulation induced on the Sagnac frequency and in particular provides a measure of low frequency effects.

Heterodyne techniques are discussed in §3.4 which are frequently used by the group to inspect high frequency signals by projecting them into the audio domain where they are more easily handled. Correlation techniques are discussed in §3.5 which allows the incorporation of prior knowledge into the detection of low level signals. The major contribution made in this chapter is a method of analysis that is entirely new to the ring group. This uses autoregressive parametric model fitting to determine the key parameters of frequency and linewidth for very short data sets. The resolution of the frequency estimates is below millihertz for data sets of only a second duration. This super-resolution is only possible by using *a priori* information about the signal. A review of such techniques is provided by Kay and Marple [64] where it is shown that using incorrect models often leads to unreliable results. For this reason the AR(2) model is chosen because it possesses all the features that the ring laser output signal is theoretically expected to contain. Both the AR(2) model and the ring output comprise a sharp spectral line with Lorentzian lineshape thus the AR(2) neatly contains all the prior knowledge available regarding the ring laser output. This model is discussed in §3.6.

3.2 Fourier Techniques

The Fourier Transform has been the cornerstone of data analysis since it was introduced by Jean-Baptiste-Joseph Fourier in 1807. With the further development of the Fast Fourier Transform by Cooley and Tukey [31] [1] the transform became

[1] A brief history of the Fourier Transform by R. N. Bracewell [24] comments that the origins of the fast method can be traced back to a paper by Gauss written in 1805. It is worth

still more popular due to its enhanced computational efficiency. In this thesis almost all references to the Fourier transform are with respect to the discrete form as this is the form most amenable to data analysis. As well as the straightforward application of the transform to data, several effects are discussed here including aliasing, zero padding and artificial line broadening. The fast algorithm itself is used more or less as a black box; only the pertinent facts regarding its performance are included here. Also considerable discussion is given to the averaging process as this has been used as standard during data collection.

3.2.1 The Periodogram and the Discrete Fourier Transform

The periodogram is a measure of the spectral content of a signal $X_1, \ldots, X_N$ and is defined for all frequencies $-\pi \leq \omega \leq \pi$ by

$$S_N(\omega) = A^2(\omega) + B^2(\omega) = \frac{2}{N}\left|\sum_{t=1}^{N} X_t e^{-i\omega t}\right|^2 \tag{3.1}$$

where[2]

$$A(\omega) = \sqrt{\frac{2}{N}}\sum_{t=1}^{N} X_t \cos \omega t \tag{3.2}$$

$$B(\omega) = \sqrt{\frac{2}{N}}\sum_{t=1}^{N} X_t \sin \omega t \; . \tag{3.3}$$

A set of discrete frequencies must be chosen for which the power spectrum is to be estimated. Any set of frequencies is allowed but usually the set of frequencies is chosen to be $\omega_n = 2\pi n/N$ where $n = 0, \ldots, N/2$. The reason for making this choice is that the basis functions form an orthonormal basis for the sample space beginning with the zero frequency term, and are complete, giving an invertible transform. The resulting spectral estimates are equal to the discrete Fourier transform and only using this choice of frequencies can the Fast Fourier transform be implemented.

Zero padding is a technique which is only mentioned for completeness. I have avoided using this technique while analysing ring laser data as it invites incorrect interpretation of spectra (see for example Kay and Marple [64]). Two basic methods are used, the first adds a string of zeros onto the end of a data set to extend the length to one that may be rapidly processed (see §3.2.2), usually the

mentioning that this work antedates Fourier's original work

[2] It needs to be noted that the FFT provided with MATLAB® is a raw transform the same as that provided in Eq. (3.2) but without the factor of $2/N$. This needs to be noted for the purposes of comparison between Fourier transform plots. An excellent power spectral estimate is built into the MATLAB® distribution in the `spectrum` routine which is based on the power spectrum estimate due to Welch [125].

final length is chosen to be a power of 2. The second adds a set number of zeros between each data point. The first of these is a useful technique when rapid processing is important but is of limited use for continuous ring laser output which is segmented into appropriate lengths before analysing. The second technique improves the appearance of the power spectral estimate by supplying additional frequency bins between those of the DFT; this aids the identification and location of spectral lines. However, zero padding does not enhance the underlying resolution of the transform made from a given finite data sequence. [75, p.43]. It might be intuitively guessed that this is the case as the process of zero padding does not add any new information to the data set.

There exist several interpretation issues regarding Fourier analysis. Those which have occurred while analysing ring output are discussed. Aliasing is one of the first problems encountered when using Fourier analysis. Consider a signal which contains a frequency component at f_0. As is depicted in Fig. 3.2 if the sample rate

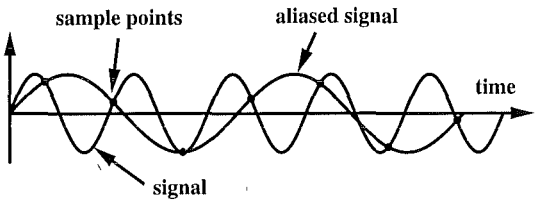


Fig. 3.2: The signal is sampled at discrete, evenly spaced time locations but the sample rate is not sufficient to identify the actual frequency of the signal. The apparent frequency is said to have been aliased down into the Nyquist window.

is not high enough a tone may not be properly represented by the sampled data. Specifically it is found that for a sampling frequency of f only frequencies from 0 to $f/2$ are properly represented; this is the sampling theorem. This upper limit is known as the Nyquist frequency and this frequency span known as the Nyquist window. A frequency component that exceeds the Nyquist limit is reflected back into the Nyquist window. If this means that the component drops below zero then it is again reflected into the window. This process continues up to the bandpass of the detector circuit. For the purposes of the ring group only frequencies up to a few hundred hertz are required to identify the audio Sagnac frequency and perhaps a few harmonics of this frequency. Unfortunately the aliasing process folds noise into the Nyquist window which degrades the signal somewhat. In fact, if the detection apparatus has an overall bandwidth B and the sampling frequency is f_s then the aliasing of high frequency white noise into the lower frequency Nyquist window gives an increase in the noise floor by a factor $2B/f_s$. Thus a low-pass antialiasing filter, which has a bandpass equal to the Nyquist frequency, is placed between the detection apparatus and the A/D converter. This mechanism is used in the SRS770 audio spectrum analyser. I have designed a suitable active antialiasing filter for use with the STROBES acquisition system which has been the major data collecting means for C-I.

Another unfortunate by-product of under-sampling a signal with a dominant tone

is that as the sample rate falls in and out of phase with the signal the apparent peak-to-peak amplitude of the signal seems to be modulated. This has created an interpretation problem with ring laser data, as a long data run requires low sample rates and the drifting Sagnac frequency means that the sample frequency and the Sagnac frequency fall in and out of phase giving the appearance of an amplitude modulation. Any genuine time dependence of the amplitude envelope has to be separated from this effect before a meaningful interpretation can be made. Consequently a Nyquist frequency is chosen to be in excess of any frequencies of interest. This limits the artifactual modulation but the mechanism is still evident for Nyquist frequencies many times the frequency of interest.

The fast Fourier transform, discussed below, necessarily requires data samples to be evenly spaced. It is possible to define a waveform better by using unevenly spaced sample points so that parts of the waveform that have greater curvature have a higher density of points. The data sets that result are invariably smaller than those required to give the same waveform definition with evenly spaced data. The phenomenon of aliasing can be almost totally removed by using unevenly spaced data points; the interpolating function far better represents the original signal. On the same theme it is worth noting that standard Fourier techniques are not well suited to analysing disjoint data sets. These typically arise when observation of a signal is intermittent for example in Astronomy (due to weather etc.) or Atmospheric Physics. However, these methods are not easily implemented. All data collected from the ring lasers are sampled at evenly spaced points in time.

The discrete and evenly spaced frequencies of the Fourier transform do not necessarily provide the best frequency resolution possible for a sampled signal. Here a signal with two coherent spectral lines (see Fig. 3.4) of different power content is considered. To emphasise the points made here the stronger tone is misaligned with the basis frequencies whereas the weaker tone is centred exactly on a frequency bin. The Fourier transform based power spectral estimate would have the weaker line rise above the stronger line as the latter has its power distributed across many bin whereas the former has all its power concentrated in a single bin. Following Priestley [96, p.399 et. seq.] it is found that the expectation value of the spectral estimate $S_N(\omega)$ is given by

$$E[S_N(\omega)] = 2{\sigma_\epsilon}^2 + \frac{1}{2N}\sum_{i=1}^{K} {A_i}^2 \left\{ \frac{\sin^2(\frac{1}{2}N(\omega+\omega_i))}{\sin^2(\frac{1}{2}(\omega+\omega_i))} + \frac{\sin^2(\frac{1}{2}N(\omega-\omega_i))}{\sin^2(\frac{1}{2}(\omega-\omega_i))} \right\} \quad (3.4)$$

for a signal which is of the form

$$X_t = \epsilon_t + \sum_{i=1}^{K} A_i \cos(\omega_i t + \phi_i) \; . \quad (3.5)$$

The signal contains K tones at frequencies $\omega_1 \ldots \omega_K$ and has an additive white Gaussian noise term ϵ_t which is stationary with variance ${\sigma_\epsilon}^2$. Notice the appear-

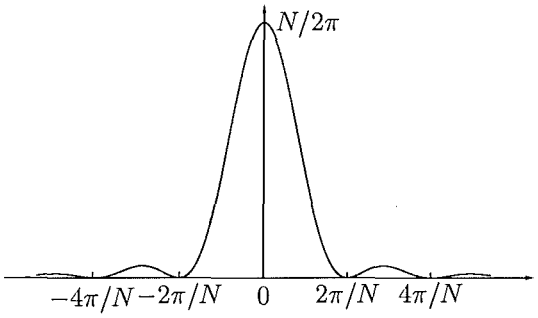


Fig. 3.3: The Fejer kernel, the contribution made to the continuous periodogram by an ideal sinusoid sampled at even intervals. Note that the zeros appear either side of the origin at intervals $2\pi/N$ whereas the value at the origin is the only nonzero value in this comb of frequencies. The number of samples in the data set is N.

ance of the expression

$$F_N(\theta) = \frac{1}{2\pi N} \frac{\sin^2(N\theta/2)}{\sin^2(\theta/2)} \tag{3.6}$$

which is known as the Fejer kernel and is found frequently in signal analysis. This

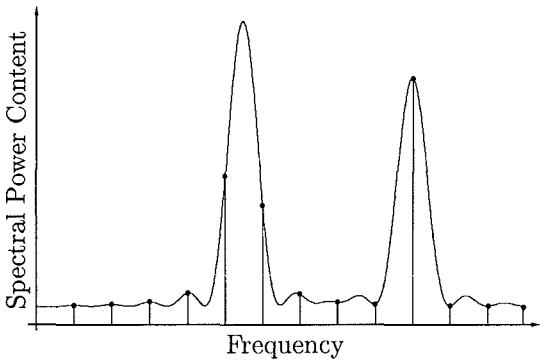


Fig. 3.4: An illustration of how discrete sampling of two sinusoids gives spread lineshapes in the frequency domain. The signal on the left is more powerful, its Fejer profile has greater amplitude, but lies between two bin so that its power is divided between several neighbouring bin. In contrast the less powerful signal on the right lies exactly on a frequency bin and all of its power is concentrated in this bin. Notice that the larger signal does not appear larger on the standard periodogram.

is illustrated in Fig. 3.3. The estimates for the power spectrum are given at the discrete locations specified by the periodogram. Under usual circumstances (with no zero padding) these sample locations are at the same spacing as the zeros in the Fejer kernel; in this case a sinusoid which coincides with a frequency bin in the periodogram will contribute only to that bin. The kernel can be used to see how the spectral estimates are found in the event that the sinusoids lie between bin as is illustrated in Fig. 3.4. Of course a finer grid of basis frequencies for the periodogram can be used which will allow greater accuracy in the location of line

centre but the effect of the corresponding Fejer kernel lineshape still limits the resolution.

A technique that is often used in conjunction with standard Fourier analysis is windowing. Windowing is a process by which a data set is altered in the time domain to achieve a particular aim usually in the frequency domain. Usually this is achieved by multiplying the sampled data stream X_j by a window F_j and finding the power spectrum of this modified time series

$$S(\omega_n) = \left| \sum_{j=1}^{N} F_j X_j e^{i\omega_n t_j} \right| . \tag{3.7}$$

One application of windowing is used to reduce artifactual line broadening which occurs when forming the Fourier transform; this already has been seen in § 3.2.1. A window such as the Hanning window may be applied to the data set before transforming and this narrows the broadened line. However, much of the analysis and processing that I have performed throughout this thesis concerns the width of a spectral line in a signal. Windowing affects the spectral linewidth and has other undesirable features such as a nonuniform transfer function. It has not been used to any significant degree in the data processing.

3.2.2 The Fast Fourier Transform

The fast Fourier transform is due to Cooley and Tukey (1965) [31] and its computational efficiency has made it the most widely used method of data analysis. Usually a data set is chosen to have a length which is a power of two but in general it is possible to consider a data set to have length N with prime factorisation ${p_1}^{n_1}{p_2}^{n_2}\ldots{p_m}^{n_m}$. It is found that the minimum number of operations required to find the fast Fourier transform [37] is $N(p_1 n_1 + p_2 n_2 + \cdots + p_m n_m)$. The algorithm is most efficient for data sets of length $N = 2^M$ where the FFT takes only $2N\log_2 N$ operations. There is no benefit in using this algorithm for a data set that has a length that is prime as this is an $O(N^2)$ process, the same as the direct Fourier transform. Modern implementations of the fast method, such as that used in MATLAB®, make use of this factorisation scheme to improve performance. As an illustration, sets of data which are of length 4095, 4096 and 4097 require a minimum number of operations to evaluate the discrete transform according to:

$4095 = 3^2{\cdot}5{\cdot}7{\cdot}13$	operations	=	126945
$4096 = 2^{12}$	operations	=	98304
$4097 = 17{\cdot}241$	operations	=	1057026

It is thus highly beneficial to select the length of a data set before performing the FFT and although a length which is a power of 2 is not always possible efforts should be made to maximise the number of prime factors in the length of the data set.

A further technique is commonly used to halve the time taken to find the transform of a real valued data set. This mechanism has been incorporated in modern implementations of the fast Fourier transform via the following formula. [94, p.398] Consider a real valued data set f_j where j varies from 0 to $N-1$. The data is repacked into a complex valued data set with length $N/2$ according to $h_j = f_{2j} + if_{2j+1}$ where $j = 0, \ldots, N/2-1$ and the FFT H_n of h_j is evaluated. The Fourier transform F_n of f_j is determined by making use of the unpacking scheme:

$$F_n = \tfrac{1}{2}\left(H_n + H^*_{N/2-n}\right) - \tfrac{i}{2}\left(H_n - H^*_{N/2-n}\right) e^{2\pi in/N} \; . \tag{3.8}$$

This process efficiently makes use of the fact that the time domain data is real and is often used in practice; for example both the SRS 770 audio spectrum analyser and MATLAB® use this technique.

3.2.3 Nonstationary Spectrum Analysis

When a discrete data set is transformed into the frequency domain a discrete set of basis functions is used to represent the data. These functions are sinusoids which run along the length of the data and as such are not well suited to handling transient effects or nonperiodic signals. In particular transient effects such as earthquakes and other impulse responses are very poorly described by steady state solutions. The spectrum of transient effects is expected to evolve with time; this information is completely lost when using the *one shot* Fourier transform.

A realization of the instantaneous spectrum may be found by using wavelets. Unlike the basis functions of Fourier analysis which extend to infinity, wavelet basis functions are only locally nonzero. The mathematics of wavelets tends to be involved; a good introduction is given by Chui [29] and another by Lang and Forinash [68]. A typical feature of wavelet analysis is that the basis is chosen to suit the source of data. For example the expansion of seismic transients in terms of Laguerre functions has been described by Båth [15, p.94] [16, p.111].

The aim of wavelet analysis is to provide information optimally about how a spectrum evolves with time; this is known as joint time-frequency analysis. This idea is used by the ring group but in a simple manner, thereby avoiding the complexities of wavelet analysis. This is possible as, in general, our spectra are only slowly varying in time. The basis that is used is simply the Fourier basis for a short local segment of data. Hence a local estimate may be provided for the spectrum by considering sections of the whole data set. One implementation breaks the data set into partitions of equal length and then plots the Fourier spectrum for each partition as a function of time (this is called waterfall in MATLAB®). An example is illustrated in Figs. 3.5 and 3.6.

Fig. 3.5 illustrates how a drifting signal appears when Fourier transformed. Very little can be deduced from this plot apart from the existence of a carrier, its harmonic and a tone at 2.5 Hz. Fig. 3.6 demonstrates the technique of estimating

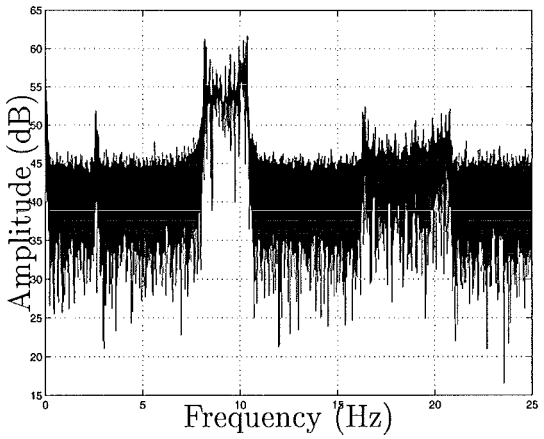


Fig. 3.5: This figure illustrates the direct Fourier transform of a 35 minute data set taken on 28 September 1995 from C-I. Although it is clear that there has been drift it is not clear how the Sagnac frequency depends on time.

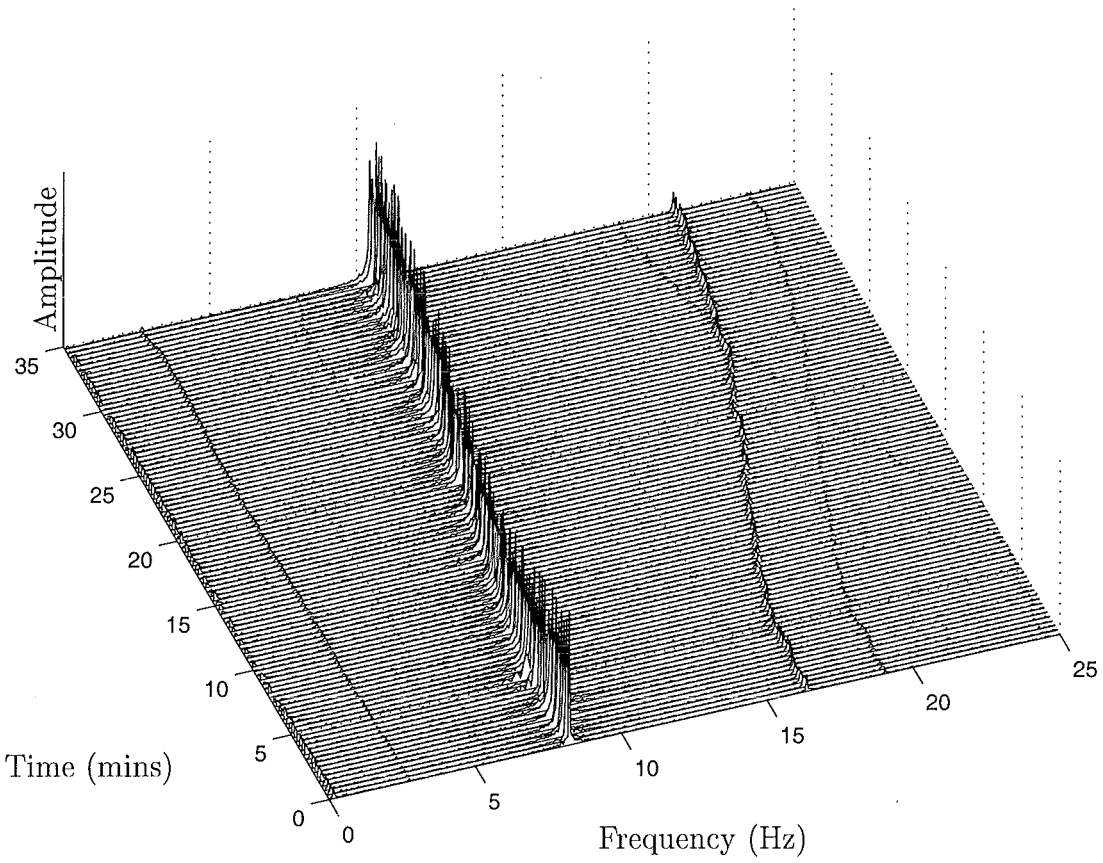


Fig. 3.6: An illustration of the time-frequency analysis used by the ring group. This waterfall plot shows how the spectrum evolves with time and uses the same data set as that used in Fig. 3.5.

the spectrum in a short period of time and then displaying its evolution as a function of time. This figure is derived from the same data set but much more information is now available. The fundamental starts at a frequency of about 58 Hz, which is doubly aliased down to 8 Hz, and drifts up to about 11 Hz. The first harmonic is also clearly seen initially at about 16 Hz and drifts by twice the

amount of the fundamental as expected. This plot allows the observation of other phenomena which are not visible in Fig. 3.5. For example the appearance of a spurious tone at about 2.5 Hz is clear in both plots but only in the second plot is this tone seen to induce a single sideband on the first harmonic and also faintly on the fundamental. This provides information about the nature of the spurious tone which may be used to isolate and remove it. Further, the second harmonic can be seen emerging from the Nyquist frequency at the start of the run; it has been reflected off the Nyquist frequency and hence drifts the opposite way to the fundamental and first harmonic.

3.2.4 The Analytic Signal

When considering a real valued signal, such as the output from C-II, it is useful to think of it as the real part of a complex signal. The usefulness of this idea lies in the process of estimating the phase of a signal in the time domain. Considering Fig. 3.7 it is easy to see how an estimate of phase may be made by eye. However, a more formal approach to estimating the phase of such data is required. This may

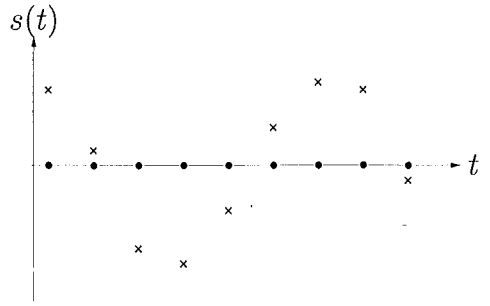


Fig. 3.7: A schematic of the data output. It is easy by eye to guess a phase at each point and hence it should be possible to write a program to do the same thing

be done by a variety of methods such as fitting expected profiles to the waveform. The general method of the analytic signal is discussed here. The concept of the analytic signal is due to Gabor [47]. The reasoning used here follows Panter [86] and especially Born and Wolf [22].

A real valued signal $s(t)$ may be expressed in terms of its Fourier spectrum according to

$$s(t) = \int_0^\infty a(f)\cos(2\pi ft - \phi(f))\,df \tag{3.9}$$

where $a(f)$ is real valued. The complex function $\psi(t)$ is associated with $s(t)$ according to

$$\psi(t) = \int_0^\infty a(f)e^{i(2\pi ft-\phi(f))}\,df \tag{3.10}$$

so that $\psi(t) = s(t) + i\sigma(t)$ where $\sigma(t)$ is defined as

$$\sigma(t) = \int_0^\infty a(f)\sin(2\pi ft - \phi(f))\,df\ . \tag{3.11}$$

It is found that $s(t)$ and $\sigma(t)$ form a Hilbert pair

$$\sigma(t) = \widehat{s(t)} = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{s(t)}{t-\tau}\,d\tau \quad \text{and} \quad s(t) = -\frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\sigma(t)}{t-\tau}\,d\tau \tag{3.12}$$

The complex function $\psi(t)$ is referred to as the analytic signal because it is analytic in the lower half of the complex plane.

An important property of the Hilbert transform is given in terms of the relation to the Fourier transform, explicitly

$$\Psi(i\omega) = \begin{cases} 2S(\omega) & \omega > 0 \\ 0 & \omega < 0 \end{cases} \tag{3.13}$$

and this provides a method for evaluating the Hilbert transform numerically. The Fourier transform of the real signal $s(t)$ is found and the negative frequency components set to zero. The inverse Fourier transform is then the Hilbert transform. This is the technique implemented in MATLAB®. The analytic signal allows estimation of the amplitude envelope, $\sqrt{s^2(t)+\sigma^2(t)}$, and the instantaneous phase, $\phi(t) = \tan^{-1}(\sigma(t)/s(t))$. From the latter an estimate of the instantaneous frequency may be found using $f(t) = \phi'(t)$.

Within the ring laser group dedrifting has been performed by using the source data to determine the current frequency and then to shift that frequency back to a predetermined frequency [111]. This is achieved by obtaining the modulation spectrum using the analytic signal technique. The drift is determined from the low frequency part of the modulation spectrum using only frequencies from 0 Hz up to a cutoff ν_c (often of the order of 1 Hz). This procedure has been successful in revealing sidebands on the carrier which would otherwise be smeared out and not observed. The sidebands are only recovered if they are outside the pass band $\pm\nu_c$. However, all the modulation information is contained in the modulation spectrum and the dedrifting procedure is not needed to recover the sidebands which appear as a single line in the modulation spectrum. The gain that dedrifting provides is to display the sidebands as two peaks either side of a carrier. This aids interpretation. Previously dedrifting has been falsely interpreted as giving a very narrow carrier signal [5,19,20,32,108,110–112] of 140 nHz width. In these instances the low frequency terms that constitute linewidth are filtered out along with those low frequency terms that comprise drift. In Chapter 4 a typical data set is analysed using several methods to determine the linewidth and good agreement is found; the linewidth for that data set is about 1 mHz.

3.2.5 The Hilbert Transform as a Digital Filter

For sample rates that are sufficiently high to avoid aliasing effects (1000 s/s for C-I circa Oct 1998) and for data sets that are sufficiently long as to allow the detection of low frequency phenomena, the end of our computing resources are soon reached. The evaluation of the analytic signal requires two Fourier transforms of

the data set. This is prohibitively time consuming for long data sets. Further, the acquisition of a data set must be complete before the Hilbert transform may be found if the Fourier method is used. An alternative method is the use of a digital filter that performs the transform. Digital filters are discussed at length by Hamming [54]. Such filters can perform many operations on evenly spaced discrete data including differentiation, integration, smoothing, frequency filtering and so on. I introduce the use of digital filters to the ring laser group to perform the Hilbert transform; this is discussed by Embree and Kimble [44, pp.202-204] and Oppenheim and Schafer [85, Chapter 7] and filter performance is characterised by Rabiner and Schafer [97]. Filters may be generated in MATLAB® using the `firls` function which creates finite-duration impulse response filters (hence FIR in the name) using least-squares error minimisation (hence LS in the name). In particular the command that generates an `N` point filter with good frequency response between `FL` and `FU` is `B=firls(N-1,[FL FU],[1 1],'hilbert')` where the final filter is `B`. I have used a 35 point filter in all of my analysis which is characterised here. The digital filter given below is convolved with the signal $s(t)$, the output being the transform $\sigma(t)$. Note that each alternate point in the filter is zero suggesting that further economising in processing is possible when the convolution is performed.

```
[0.038135 0 0.024179 0 0.032403 0 0.043301 0 0.058420 0 ...
0.081119 0 0.120167 0 0.207859 0 0.635163 0 -0.635163 0 ...
-0.207859 0 -0.120167 0 -0.081119 0 -0.058420 0 -0.043301 ...
0 -0.032403 0 -0.024179 0 -0.038135]
```

The transfer function of the filter, given in Fig. 3.8, shows that the filter also acts as a very broadband filter from about 0.02 to 0.48 of the sample frequency. Between these frequencies the filter is very close to unity, varying by about 0.5 dB. This attenuation does not significantly affect the Hilbert transform provided the carrier frequency does not move to within 0.03 (multiplied by the sampling frequency) of either end of the Nyquist window. The raw data is usually band passed before the filter is applied. This is necessary as the digital filter, which is differential in nature, does not work well if the signal has a d.c. component. Also the band pass process can be achieved by use of a digital filter; further economy is achieved when applying two successive digital filters by convolving the two filters first and then applying just the one filter. Thus a filter is produced that bandpasses and finds the analytic signal at the same time.

The digital filter approach has benefits over the Fourier method. The Fourier method uses the data from the entire sample to estimate the phase at one location. This is not realistic in the event that the carrier is stochastically drifting; samples that are further spaced apart in time become increasingly independent of each other. The digital filter method is localised, deriving its estimate from only a relatively few samples either side of its centre. This technique has been tested on many data sets and, provided the carrier frequency does not drift to the edges of the Nyquist window, the results are unchanged from the analysis which uses

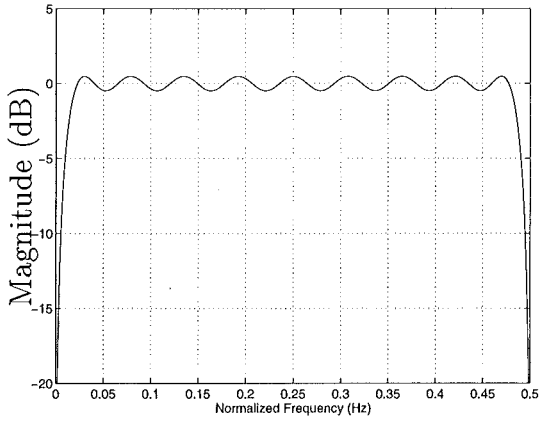


Fig. 3.8: The Transfer function of the 35 point Hilbert filter. Note once again the standard of normalising the responses of filters etc. in the frequency domain to a Nyquist window 0.5 Hz wide.

Fourier transforms to form the analytic signal.

The digital filter approach allows the analytic signal to be calculated and saved as data is being acquired, using the fact that at 400 samples per second the computer lies idle for most of the time between samples. Thus the bold claim is made that the analytic signal takes zero time to calculate as it is available as soon as the data acquisition run is complete. More precisely 18 floating point operations (flops) are required to evaluate the complex analytic signal value for each sample. With the sampling rate at 400 samples per second the computer has to evaluate about 8000 flops per second which requires less than 0.1% of the CPU time of the PENTIUM® series processors from INTEL®.

3.3 Spectrum Averaging Techniques

Averaging techniques improve ease with which signals may be isolated from broadband noise. In this section the nature of two quite different averaging methods, rms and vector, is explored. In particular their effects on the statistics of spectral estimates are explored. In all cases when evaluating the average of n spectra the complex spectral estimates for each frequency bin are evaluated in turn. The n estimates for that bin are collected, the average performed and the result placed in the appropriate bin in the resulting averaged spectrum. The results given here do not consider non-stationary spectra, for example those with a drifting signal.

The first technique considered is rms averaging which takes the direct average of the estimates for each bin of the power spectrum. This method is described by Welch [125] where it is shown that the optimum value of overlap for data segments is zero for an indefinitely long stream of data or 50% for a data stream of finite known length. In rms averaging the rms of the absolute value of the n estimates is formed for a particular bin. This has the effect of reducing the variance of the spectral estimate while the estimate itself tends to a finite nonzero value. Any small amplitude coherent feature becomes apparent as more spectra are included

in the average so as to smooth the noise spectrum.

The other method is vector averaging which finds the average of the n complex estimates. Unlike rms averaging, vector averaging allows incoherent effects (noise) to cancel. A qualitative description of this cancellation process would suggest that only a coherent signal will not be averaged to zero. However, for this to be true data collection for each spectrum must be initiated in phase synchrony with the signal. This requires synchronous detection for an entire spectrum. Such a detection method is used when a signal is dependent on an external modulating effect. The modulation can be used to provide a trigger for data collection to start. Hence any coherent effect arising from the modulation will add constructively under vector averaging.

A comment should be made at this point regarding the use of a photomultiplier tube as a detector. The current from the device is linearly proportional to the intensity of the light that is incident on it. This current is passed to a transimpedance amplifier and the resulting voltage is appropriately filtered and fed to the A/D converter. Hence, the unusual situation where a recorded voltage is linearly proportional to power instead of quadratically proportional arises. Normally the power spectral density is estimated by the square of the amplitude of the Fourier transform but here the proper power spectral estimate uses the absolute value of the Fourier transform. Superficially this only has the effect of dividing the power spectral estimates by 2 (on a logarithmic scale) but when the statistics of the spectral estimates are considered the difference is more significant. In the next section the power spectral estimates for a white noise signal are found to be given by the Rayleigh distribution

$$P(r) = 2kre^{-kr^2} \tag{3.14}$$

where r is the amplitude of the spectral estimate. This estimate is usually given by the modulus square of the Fourier component which is governed by the chi-squared distribution on 2 degrees of freedom (for example see Priestley [96, p.397])

$$P(r) = \begin{cases} \frac{1}{2}e^{-r/2} & r \geq 0 \\ 0 & r < 0 \end{cases} . \tag{3.15}$$

3.3.1 Effects of Averaging on White Noise

The estimate of the complex Fourier component of a white noise process at a given frequency is governed by the probability distribution, $P(r, \theta)$ where r and θ map the complex plane. This distribution is the same for all frequencies hence the use of the term *white* in analogy with the uniform optical spectrum of white light. In the Ring Laser Group this has been described as grass but I prefer the more widely used expression, white noise spectrum. A white noise data sequence comprises independent identically distributed numbers. In forming the Fourier transform each component in the transform is the weighted sum of these random numbers. An extension of the central limit theorem shows us that no matter

what the distribution of the time domain numbers the probability distribution of the components in the frequency domain are Gaussian over the complex plane centred at the origin. Hence the probability that a particular spectral estimate is in the area $rd\theta dr$ is

$$P(r,\theta)rd\theta dr = \frac{k}{\pi}e^{-kr^2}rd\theta dr \tag{3.16}$$

where k defines the spread of the distribution and the metric term $rd\theta dr$ is usually not included. As has been pointed out earlier (§3.3) the absolute value of the spectral estimate and not its square gives the estimate of the power spectrum. To obtain the probability distribution of these absolute values the θ-dependence needs to be removed according to

$$P_1(r) = \int_0^{2\pi} P(r,\theta)rd\theta = 2kre^{-kr^2} \quad r > 0 \tag{3.17}$$

where $P_1(r)$ denotes the probability distribution for the estimate arising from only one data set. In Appendix B it is shown that the probability distribution of the rms average of n numbers chosen from $P_1(r)$ is given by

$$P_n(r) = \frac{2(nk)^n}{(n-1)!}r^{2n-1}e^{-nkr^2}\,. \tag{3.18}$$

The appendix provides a proof of this result which has not been found in the literature. In Fig. 3.9 the initial probability distribution, which is re^{-r^2}, becomes

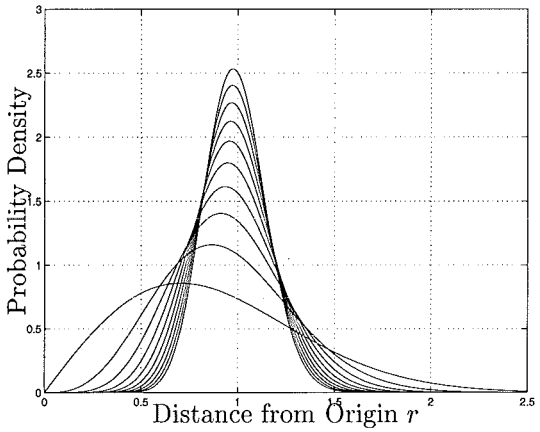


Fig. 3.9: Plots of $P_n(r)$ for $k = 1$ and $n = 1,\ldots,10$. As progressively more spectral estimates are made the number that appears in the averaged spectrum appears as if chosen from a distribution that narrows as $1/\sqrt{n}$.

a Gaussian for large n, which follows from the central limit theorem, and that the width of the Gaussian goes down as $n^{-1/2}$. For completeness the mean value $\bar{r}_n$ of the distributions is also evaluated

$$\begin{aligned}\bar{r}_n &= \int_0^{\infty} rP_n(r)dr \\ &= \sqrt{\frac{n\pi}{k}}\frac{(2n)!}{(n!2^n)^2}\end{aligned} \tag{3.19}$$

and in the limit as $n \to \infty$, $\bar{r}_n \to 1/\sqrt{k}$.

Attention is now turned to vector averaging. For a single Fourier transform each spectral estimate is chosen from the distribution given in Eq. (3.16). The average of the complex values in a bin is found and then the absolute value of the result is plotted. The probability that the number lies between r and $r+\mathrm{d}r$ is given in Eq. (3.17). The probability distribution for the average of n complex numbers chosen from $P_1(r,\theta)$ is

$$P_n(r,\theta) = \frac{nk}{\pi} e^{-nkr^2} \tag{3.20}$$

which is Eq. (3.16) but with k replaced by nk. The probability distribution for the absolute value of such an estimate is given by

$$P_n(r) = \int_0^{2\pi} P_n(r,\theta) r d\theta = 2nkre^{-nkr^2} \tag{3.21}$$

The validity of this result may be checked by performing an inductive proof. It is proposed that the distribution that governs the sum of n complex numbers chosen from $P_1(r)$ is $P_n(r)$. The proof requires us to demonstrate that the distribution that governs the sum of a number from each of these distributions is $P_{n+1}(r)$. Such a result has been verified but it is not enlightening to include the calculations here. The average value of these distributions is found and obeys the following limit:

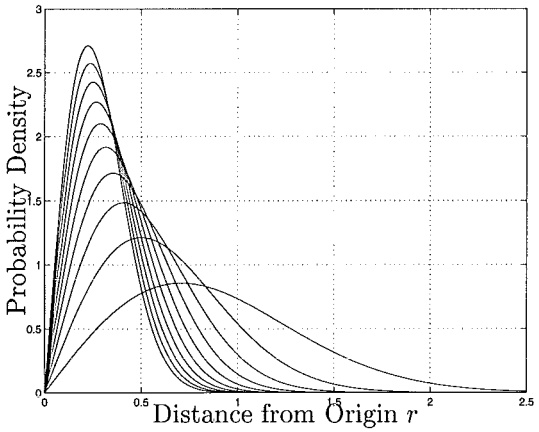


Fig. 3.10: Plots of $P_n(r)$ for $k = 1$ and $n = 1, \ldots, 10$. Increasing n has the effect of narrowing the distribution while the expectation value tends to zero.

$$\bar{r}_n = \sqrt{\frac{\pi}{4nk}} \to 0 \quad \text{as} \quad n \to \infty \,. \tag{3.22}$$

The appearance of these distributions is illustrated in Fig. 3.10.

In summary, the rms average narrows the width of the probability distribution that governs the spectral content of each bin with the mean of this distribution tending toward a finite nonzero limit. The vector average also squeezes the distribution but the mean tends towards zero and hence noise may be reduced without limit with enough averages.

3.3.2 Externally Locked Ring Laser Operation

In Chapter 5 a locking phenomenon is described. One of the ring laser mirrors is sinusoidally oscillated using a piezo and it is found that when the frequency of this modulation approaches the Sagnac frequency the two lock together. This provides an alternative mode of operation and an excellent example of how averaging techniques allow small signals to be isolated amidst noise. While locked to an external signal a Lissajous figure is formed on an oscilloscope with the ring laser output and the external signal controlling the x-y oscilloscope display. A variety of rudimentary tests show that the laser is still highly sensitive to external effects. Small shocks (stamping feet) have the effect of jittering the Lissajous pattern while more gentle flexing of the base granite (pushing with finger) shifts the phase and amplitude of the pattern. Thus the locked carrier frequency is sensitive in a phase modulation and amplitude modulation sense. The main benefit is that spectra may by synchronised with the piezo driver allowing rms averaging. This has been done and the results are shown in Fig. 3.11. The uncertainty in the individual spectral estimates is sufficiently low that small amplitude spectral features may be detected. The raising of the noise floor either side of the carrier frequency from 0.1-0.3 Hz indicates that some low frequency spectrally broad phenomenon is amplitude/phase modulating the ring laser. The coupling to the laser is most likely through the bed rock or acoustic. Possible candidates include electric motors (fans), wind vibration, micro-seisms, traffic or even crashing waves on the coast. Unfortunately no firm conclusions can be drawn from this result as the cavern was not acoustically quiet at the time that the measurements were taken.

The results displayed in Fig. 3.11 may be used to estimate the magnitude of the modulation that induces the sidebands. The noise floor is located at -48 $\mathrm{dBV_{rms}}$ with the raised region between 0.1 and 0.3 Hz at approximately -46 $\mathrm{dBV_{rms}}$; the carrier frequency is 81.5 Hz and has amplitude 12 $\mathrm{dBV_{rms}}$. Thus the power density of the sideband signal is about 1.6 that of the background noise. The following calculations estimate the magnitude of the rotational modulation, Δf, required to induce sidebands with this power content.

In order to estimate the size of the effect that induces this sideband the broad sideband is modelled by a narrow sideband with modulation frequency f_m =0.2 Hz. The power of this narrow sideband has to be scaled so that it equates to the same total modulation strength. The broad feature is spread over 0.2 Hz so that with a resolution bandwidth of 7.63 mHz the new sideband is required to be 26 times (14 dB) the amplitude. The equivalent sideband therefore requires -32 $\mathrm{dBV_{rms}}$ amplitude in order to represent the same magnitude of modulation. The relative heights of a carrier and its first sideband are separated by N $\mathrm{dBV_{rms}}$ where

$$N = 10\log_{10}\left(\frac{J_0(\beta)}{J_1(\beta)}\right) \approx 3 - 10\log_{10}\beta \tag{3.23}$$

and β is the modulation depth $\Delta f/f_m$. Taking $N = 44$ dBV gives $\beta = 8 \times 10^{-5}$ and hence the size of the frequency excursions is $\Delta f = 16\,\mu$Hz. This represents a

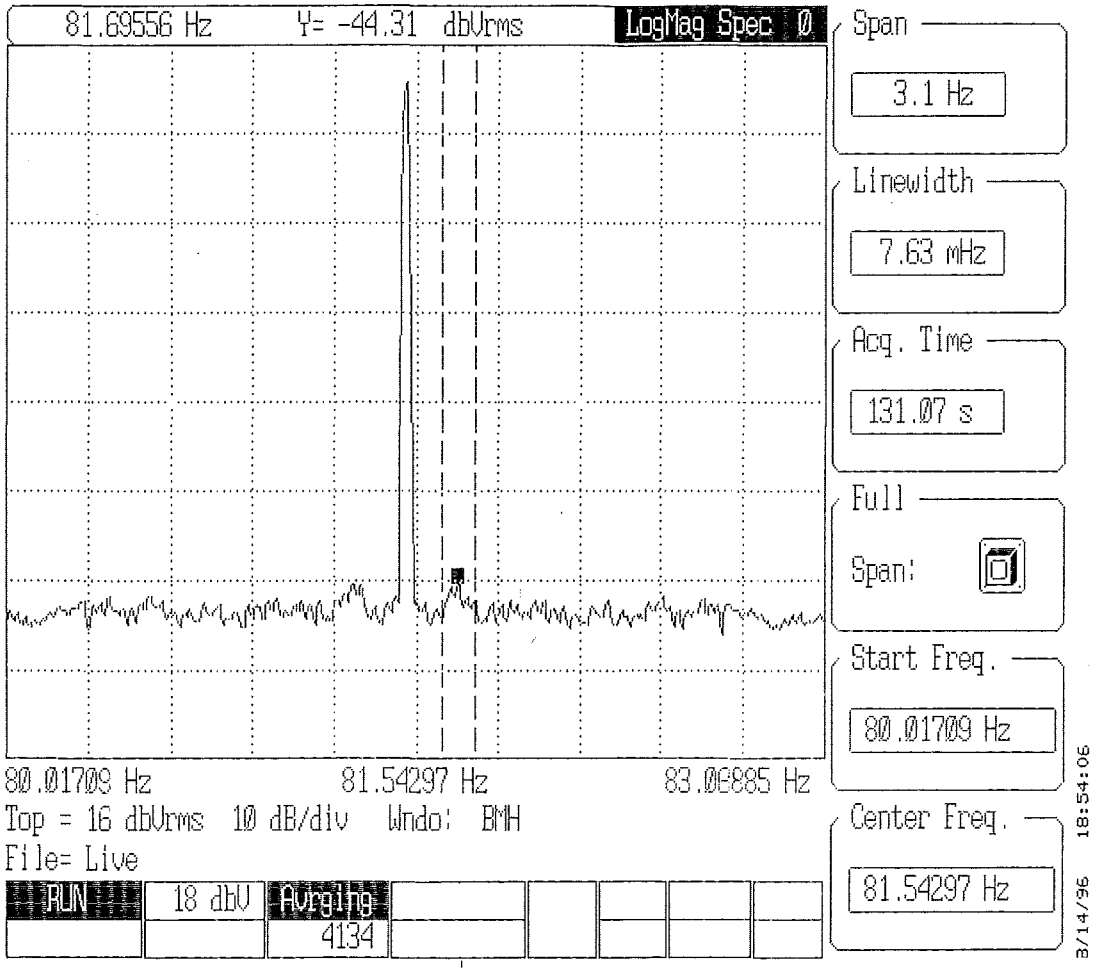


Fig. 3.11: A printout from the SRS770 audio spectrum analyser is shown which illustrates the detection of sideband structure about a stable carrier signal from C-I. Several thousand rms averages were required to reveal this phenomenon. Data collection was ceased when the Earth induced beat frequency drifted sufficiently far from the piezo modulation frequency that the two were no longer locked together.

modulation rate which, expressed in terms of the fraction of the Earth's rotation Ω_E, is $2 \times 10^{-7}\Omega_E$.

Frequency locking offers a new mode of operation for ring laser gyros. The high sensitivity is demonstrated here but unfortunately no firm conclusions can reasonably be made from these results. C-I is highly sensitive to very small environmental effects such as audio vibrations and as such the phenomenon observed here cannot be isolated from artifactual phenomena. C-II is designed to be far less sensitive to these spurious signals and hence would be a better instrument to test this style of operation. An investigation into the use of frequency locking for detection of small signals is a subject for another thesis.

3.4 Heterodyne and Super-heterodyne

Heterodyning is in principle exactly the same as amplitude modulation (AM) (see the Radio Amateur's Handbook [69, p.58 and p.244]). A received signal has its amplitude modulated by a local oscillator and sum and difference frequencies are produced in the process. The local oscillator generates AM sidebands on the signal; in this manner a duplicate of the spectrum may be shifted to any convenient place in the spectrum. A super-heterodyne detector is one which heterodynes a carrier down to a specific frequency for which a receiver circuit can be tailored (see Horowitz and Hill [59, p.895]). The tailored circuit only operates at one internal frequency. This allows superior handling of the signal through the use of selected components such as crystal filters. A problem with heterodyning is that spurious images are commonplace. For example if a carrier at 80.1 MHz is modulated by a local oscillator at 80 MHz then a 100 kHz signal is observed. Any additional noise component contained in the source at 79.9 MHz is also heterodyned to -100 kHz and is added to this 100 kHz signal. Suitable filtering before the heterodyne section can eliminate this problem.

The use of a secondary heterodyne is commonplace. This comprises an internal oscillator to heterodyne the internal operating frequency of the detector down to an audio tone. This is known as a beat frequency oscillator (BFO). Such a commercial device is used to detect the presence of a free spectral range signal from the laser considerably better than simply waiting for the *quiet* expected in the receiver due to the presence of a carrier.

If the frequency of the carrier is known (perhaps due to a pilot frequency broadcast) then this can be used as the local oscillator and the heterodyned output is passed through a low pass filter to give a measure of the instantaneous amplitude of the signal; hence this is an AM detector. This is known as synchronous detection (or *homodyne* detection or *phase sensitive* detection). If the carrier frequency is not known then the signal can be used in conjunction with a phased locked loop to regenerate the frequency and then the synchronous detection employed [59, p.652].

Heterodyne techniques have been exploited by the ring group in several areas. One of the primary roles of a heterodyne method is to detect the presence of a tone at the free spectral range frequency of the ring laser. In C-I this is $\approx$86 MHz and in C-II $\approx$75 MHz. The presence of a tone at the free spectral range frequency indicates the presence of more than a single longitudinal mode. Initially this detection was achieved using a hand held rf receiver connected to the photomultiplier output which is tuned in steps of 5 kHz. The presence of the tone is indicated by silence from the receiver or a tone if the BFO is switched on. A more sensitive method injects the PMT output into a mixing element along with a known frequency from a Hewlett-Packard rf signal generator (HP 8648A) which is locked to the 10 MHz signal provided by the ring laser GPS system. The Intermediate frequency (IF) is inspected using the Stanford Research Systems audio spectrum analyser (SRS 770) which gives a spectrum from 0-100 kHz. This

layout is depicted in Fig. 3.12. The spectrum analyser has many built-in features including several types of averaging. In this case either exponential or rms averaging is used to determine the presence of a faint tone. This use of averaging in conjunction with a visual presentation proves superior to the use of the BFO in conjunction with audio presentation and it has been demonstrated that a tone that would be missed by the rf receiver is visible on the analyser.

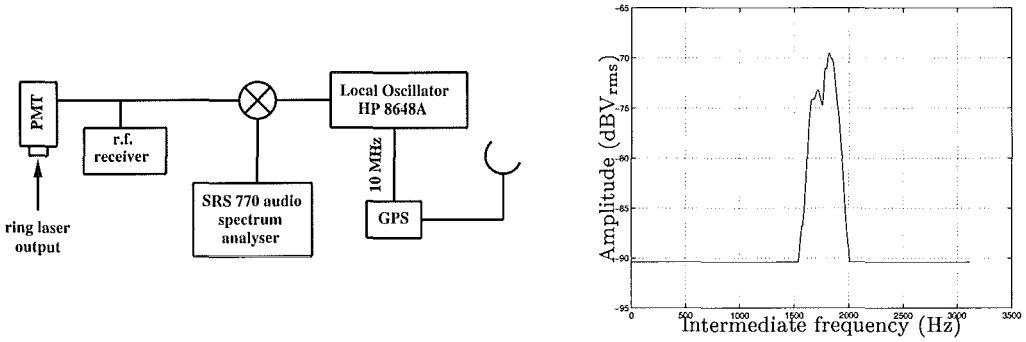


Fig. 3.12: On the left is a schematic of a heterodyne system. On the right is the output of the free spectral range signal of C-II which has line centre at about 74.97153 MHz in data recorded on 3 March 1997. The line appears to have features suggesting dispersion.

The job performed by the arrangement in Fig. 3.12 could of course be performed by an rf spectrum analyser. Unfortunately other equipment has been given higher priority for purchase up to the time that these comments were written. The apparatus shown is thus ad hoc but does have certain benefits. Only the most expensive rf spectrum analysers have the resolution that is provided by the SRS 770; this device allows windows which are sub-hertz in width, each window comprising 500 discrete frequencies. As an example of the technique Fig. 3.12 also illustrates a typical fsr lineshape when several longitudinal modes are oscillating. There appear to be two main lobes. This may be due to the presence of multiple fsr tones separated by dispersion and superimposed to form the complex lineshape seen. In this event there may be as many as five tones present.

3.5 Correlation Techniques

In addition to the methods that have now been introduced and used to analyse the current ring laser output I propose here a collection of methods that should be considered for use in experiments yet to be performed on the ring laser. The experiments envisaged involve the modulation of some parameter and the identification of this modulation in the ring laser output. Examples of the sorts of parameter that may be modulated are the electric and magnetic fields on the research arm, the local gravitational field, the intensity of a separate beam that passes through the gain section and many others. The modulation of the parameter need not be regular. Methods that could be used to retrieve an irregular

signal from a data stream are considered here; these techniques are called correlation techniques. Correlation is used in a great many areas of signal broadcasting, detection and enhancement. The principle is straightforward. If a signal is suspected to contain a feature that is known to the observer then this may be used to aid detection. The Fourier transform is particularly suited to detect signals that have the feature of being periodic. The discussion here is mainly restricted to the process of detecting signals with features other than periodicity.

The ring laser is potentially an extremely sensitive detector of non-reciprocal effects which may be induced in the lasing cavity. A scheme is discussed by Cooper and Stedman [32] in which an external magnetic field interacts with the lasing mode inside a ring cavity. It is proposed that the coupling between the magnetic field and the cavity photons is due to both axion and QED induced vacuum birefringence. The proposed experiment aims to place limits on the axion mass via a null result. The magnetic field is modulated and hence the non-reciprocal effect on the path length is also modulated. If the effect is large enough the induced sidebands on the Sagnac induced output carrier will become visible. The modulation scheme proposed is sinusoidal but need not be. The rate at which the magnetic field can be modulated is necessarily low for strong continuous fields. Low frequency modulation may well be lost in the $1/f$ frequency fluctuations induced by the active gain medium. However, if the modulation scheme is not periodic in nature then there is little chance of the gain medium fluctuations generating just the correct changes in frequency to match the predetermined pseudo-random code. The detection scheme would not be limited by $1/f$ noise.

In all experiments proposed for the ring laser that require an induced effect on the working parameters, particularly ring laser perimeter, this kind of alternative modulation ought to be considered. In this section two kinds of pseudo-random modulation are discussed. The first is periodic which repeats a section of code and the second is a single shot process. Either of these are workable as modulation schemes for the ring laser.

The basic ideas of correlation are as follows. The correlation between two functions $f_1(t)$ and $f_2(t)$ is defined as

$$C_{12}(\tau) = \int_{\infty}^{\infty} f_1(t) f_2(t+\tau) dt \tag{3.24}$$

and the autocorrelation function $R(\tau)$ is similarly defined by setting $f_2(t) = f_1(t)$. Note that the autocorrelation function is necessarily an even function. The correlation function has a natural extension in the discrete time basis

$$C_{12}(\tau) = \frac{1}{N-\tau} \sum_{t=1}^{N-\tau} f_1(t) f_2(t+\tau) \tag{3.25}$$

where $f_1(t)$ and $f_2(t)$ are defined at the points $t = 0, 1, \ldots, N$ and $\tau = 0, \ldots, m$.

The Wiener-Khintchine theorem is a very wide ranging theorem which is better known in a simplified form. In this form the theorem states that the autocorrelation function of a zero mean stationary random process $X(t)$ has a power spectral

density function $h(\omega)$ which is the Fourier transform of the autocorrelation function $R(\tau)$

$$h(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega t} R(\tau) d\tau \tag{3.26}$$

The more general theorem relates to the existence of the Fourier transform and of the autocorrelation function but these concerns are not important in the discrete time case. This is sometimes referred to as the indirect method of arriving at the power spectral density and is used in the implementation of the Maximum Entropy spectral analysis, model selection in random process analysis and the detection of weak signals in noise.

3.5.1 Using *a priori* Information

An application for correlation techniques exists in optimal signal detection. Suppose the signal is known and all that is required is knowledge of how much signal is present, if any at all. An example of this is given in § 3.5.3. This method is often discussed in terms of matched filters where the implementation is often easier using correlation in the time domain rather than filters in the frequency domain. Here a correlation detector is made according to Fig. 3.13 where the expected signal $g(t)$ is known.

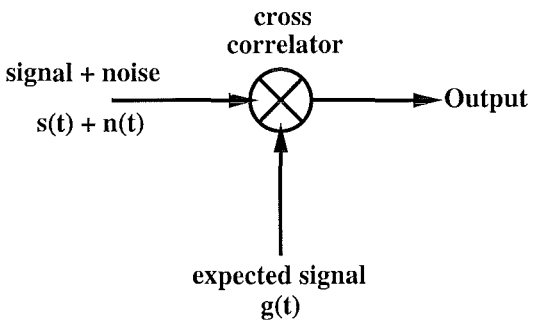


Fig. 3.13: Block diagram for a correlation detector.

Now the matched filter method involves designing a filter with transfer function $H(\omega)$ which is optimised to detect the frequency characteristics of $g(t)$. It is possible to show that the matched filter and the correlation detector are mathematically equivalent. [30, Appendix F]

One feature of correlation detection that is quite out of the reach of Fourier techniques is use of non-periodic structure in signals. Whereas the Fourier idea is to use a periodicity to add signal constructively while noise cancels, the correlation detector does not require events to be evenly spaced in order for them to add constructively. The chances of receiving a false pseudo-random signal by accident are much lower than for a false coherent signal as coherent signals are commonplace. Pseudo-random signals are used widely in atmospheric research

and especially radar where high amplitude signals are transmitted and echos are searched for in the return signal. Often echos are of very low amplitude relative to the background noise. This problem is overcome to a great extent by using a pseudo-random code in the transmitted signal. The echos are then known to have a particular pseudo-random structure which can be searched for in the return signal using a correlation detector.

The code that is used may be repeated periodically and the return signal detected using an autocorrelation detector with variable delay. These types of systems are implemented using feedback shift registers [114, pp.318]. Alternatively a single shot of pseudo-random code can be transmitted. These codes are usually Barker codes [12, 114, 122] which have the property that their autocorrelation functions have minimal side-lobe structure and maximal centre spike. There are only a finite number of Barker codes which are presented in Table 3.5.1. The simplest

N	Sequences
2	+ +, + -
3	+ + -
4	+ + - + , + + + -
5	+ + + - +
7	+ + + - - + -
11	+ + + - - - + + - + - +
13	+ + + + + - - + + - + - +

Table 3.1: It has been proven that these are the only Barker code sequences. Naturally, all the above sequences may be reversed or have polarity changed and still satisfy the Barker criterion [114, 122].

Barker code uses two evenly spaced pulses as the signal. This is the basis of the example discussed in § 3.5.3. If the code is thought of as a vector (1,1) then the cross correlation function is (1,2,1) suggesting that the side-lobes are half the height of the central lobe. A Barker code of length N gives a maximum lobe height of 1 and central spike of height N. The code that is most widely used in the rf communications field is the 11-bit Barker code.

3.5.2 Waveform Reconstruction

One further benefit of correlation which is relevant to the Ring Laser Group is waveform reconstruction (implemented using a boxcar integrator). A periodic waveform may have superimposed broadband noise which obscures the precise features of the underlying waveform. A method is described here for reducing this noise, theoretically without limit, allowing precise knowledge of the waveform. This is of particular relevance to the work reported in Chapter 5 where a ring laser model is fitted to the ring laser output waveform. More precise knowledge of this waveform improves the reliability of the fitting process.

If a data stream is known to have a periodic signal content and its precise waveform is required then the detector shown in Fig. 3.14 may be used to recover this waveform. The signal is cross correlated with a series of pulses which are

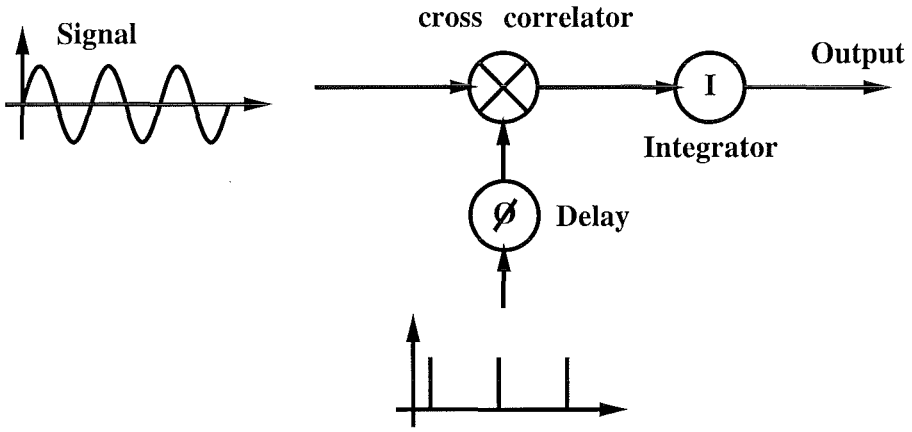


Fig. 3.14: The block diagram for a signal reconstructor.

spaced at the known period of the signal being probed. The output is passed to an integrator which constructively adds the correlation with the signal but the noise component adds destructively and becomes an ever-decreasing fraction of the total. Theoretical noise-free reproduction of the signal wave form is possible but in practice the signal to noise ratio will be enhanced by a factor dependent on the run time or the integrating time constant of the integrator. The process only gives the waveform at a particular phase; hence the relative phase of the input signal and the pulse series has to be varied through 2π in order to give a complete cycle of the waveform.

3.5.3 Lunar Ranging Data Signal Enhancement

Two illustrative examples of the successful application of correlation techniques are provided here. The first example shows how correlation methods may be used to find the ultimate resolving power of a radar transmitter/receiver. This example serves the purpose of demonstrating that such techniques achieve great signal enhancement. The second example looks at lunar ranging data. Prior knowledge of the signal is used along with a numerical correlator (in post-processing) to increase the signal to noise ratio of the data.

In a paper on radar meteor observations by Ellyett and Fraser (1955) [43] a technique is described for measuring directly the ultimate performance of a radar transmitter/receiver array. The array transmits short rf pulses and the receiver detects any reflected radiation, the source of the reflection being the ionised trail of a meteor entering the atmosphere. In order to measure directly the minimum detectable power a circuit was used to inject two closely spaced pulses of equal but variable power into the detection circuit in the place of the antenna.

Many different pairs of pulses were input of various power and their arrival times and power were noted for later reference. The detection process involved the exposure of photographic film by a beam which has an intensity proportional to the detected power. The film was developed and distributed to colleagues who simply noted where they were sure they could see a pair of dots on the image. A remarkable agreement was found. The minimum power could be determined by noting where the cutoff threshold occurred; this was quite distinct, below a certain power no pairs were seen, above another power all pairs were seen and the indeterminate band was narrow (see Fig. 3.15). The important feature of

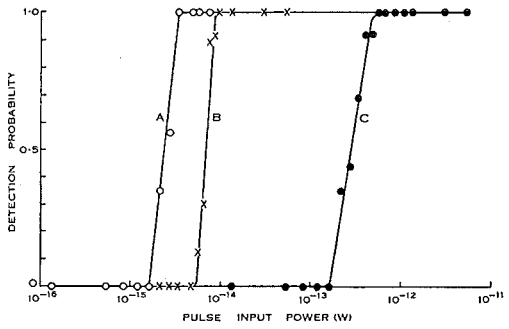


Fig. 3.15: Curve A relates to a pulse width of 26 μs, curve B to 8 μs and curve C to 3 μs.

this method of detection is that the presence of a single dot on the image is not interpreted as an echo. The likelihood of a pair of dots occurring with the correct spacing due to the noise is extremely low. The *a priori* knowledge that a pair of dots is expected is neatly incorporated into the detection scheme. In one sense the process can be thought of as a practical implementation of a deconvolution.

In order to automate the procedure described above it is necessary to decide what the essential features are. The first step that is done is to identify those parts of the image that are dots. Next, the observer has to decide whether two dots occur next to each other at the appropriate spacing τ. The first of these steps is a correlation function, in the time domain this is the cross correlation between the time domain data stream and a pulse of known form. The second of the steps uses the cross correlation stream and decides whether two pulses are present and τ apart. One implementation divides the cross correlation stream into two, one being delayed by τ, and takes the product of the two resulting data streams. This has the desired feature that if a dot occurs on its own then it will produce a pulse at one point in time t but no pulse at $t \pm \tau$, the product will be much lower than if there were two dots. Fig. 3.16 depicts a simplified version of the process. The spacing of the signal points is an integer number of sample bin which differs from the data that is considered next.

The Forschungsgruppe Satellitengeodäsie has, as one of their projects, a lunar and satellite ranging system. The essential idea is to direct a pulsed Nd:YAG-Laser into a 75 cm telescope and aim the beam at a satellite. The laser produces a series of Gaussian pulses which increase linearly in intensity to a maximum and then decrease in the same manner. A Pockels-cell restricts the exiting profile

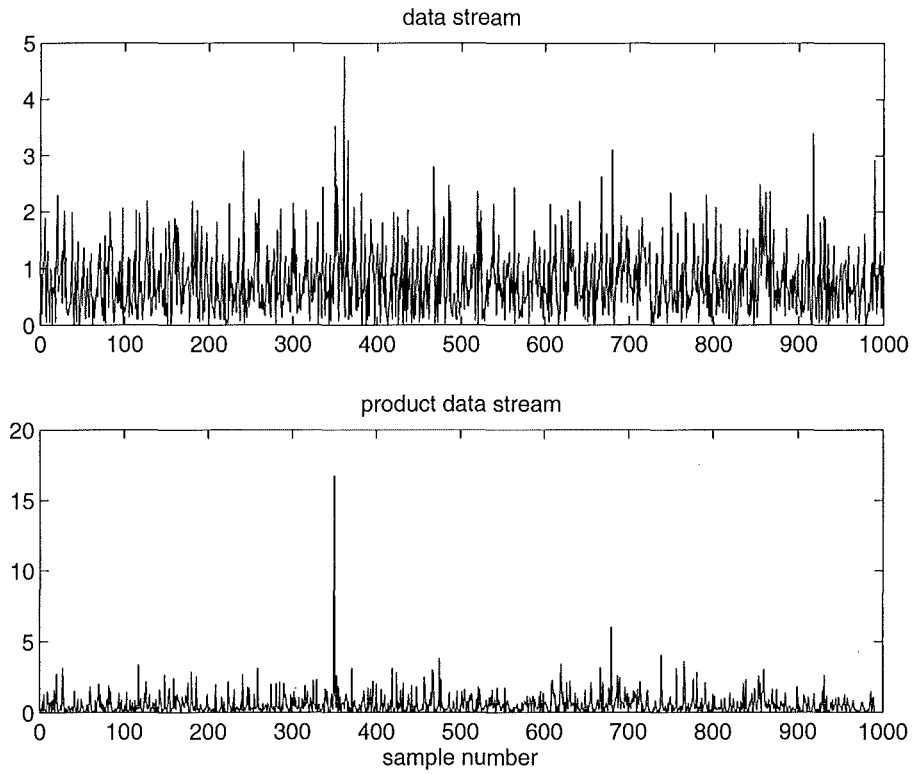


Fig. 3.16: The upper plot depicts a data stream which consists of the modulus of 1000 normally distributed numbers, $N(0,1)$, and a 0 dB signal added at the 350th and 360th sample. Once it is known that the signal is supposed to contain two pulses spaced 10 samples apart then the correlation search can be used and gives the lower plot. The observed pulse is approximately 10 dB clear of the noise floor.

to just the maximum pulse plus the subsequent pulses. The pulses in the data set under consideration are spaced by 7 ns and have width 200 ps. Information about the precise appearance of the return signal is thus available. Unfortunately the intensity of the return pulse is so low that only a few photon detections are encountered per pulse set. The upper plot in Fig. 3.17 shows a set of data where the multiple reflected pulses can be readily seen.

With the use of the correlation ideas outlined here the appearance of the output data stream is improved by using our knowledge of the expected output profile. The lower plot of Fig. 3.17 shows the result of applying the algorithm to the given data. The improvement in the visibility of the signal is from about 4 dB to about 10 dB.

Certain points of the method used here need explanation. The spacing of the bin is 1.2 ns whereas the pulses are separated by 7 ns; not an integral number of bin. It can be seen in the raw data that spacing between the first two spikes seems to be 8 bin whereas the next two spikes are separated by 9 bin. This is due to the spillover from one bin to the next. If the 200 ps wide pulse lies between two bin its contribution to the total signal is divided in two. For this reason the profile

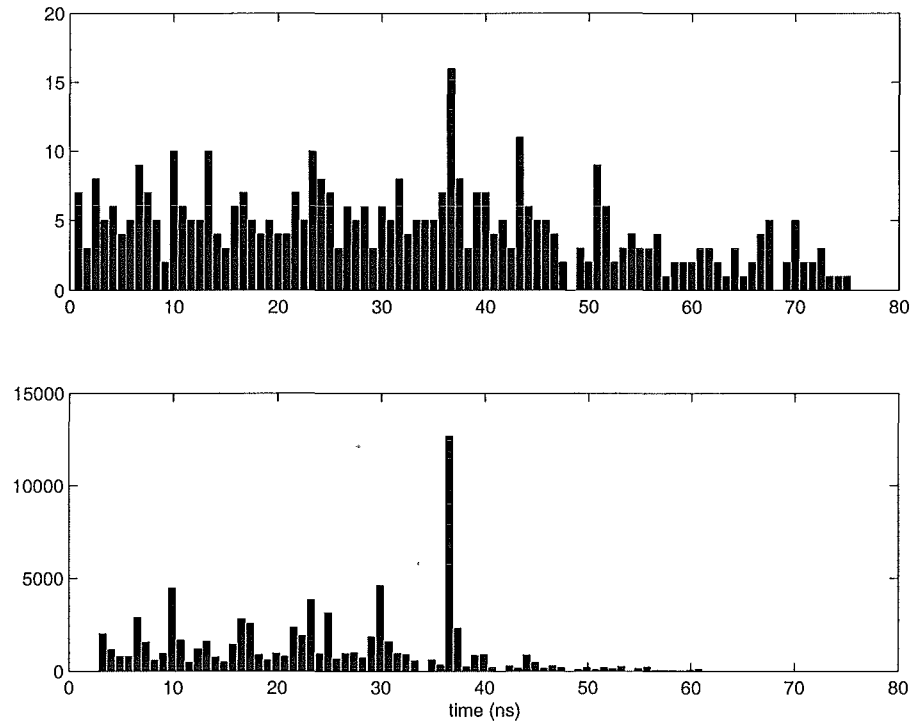


Fig. 3.17: A set of counts for reflection off Apollo 15. Upon close inspection it is possible to see that side lobes have appeared either side of the major return pulse. The major return pulse occurs at about 36.58±0.17 ns (±5 cm) after the initial sample.

of the cross correlation signal is allowed to vary its position relative to the centre of the nearest bin.

As mentioned earlier this signal recovery could be further enhanced by the use of pseudo-random code such as the Barker code. However, in the case of lunar ranging the rate at which pulses occur is considerable and hence the means of implementing coding is difficult. If the pulse rate could be made slower then such an implementation could be effected.

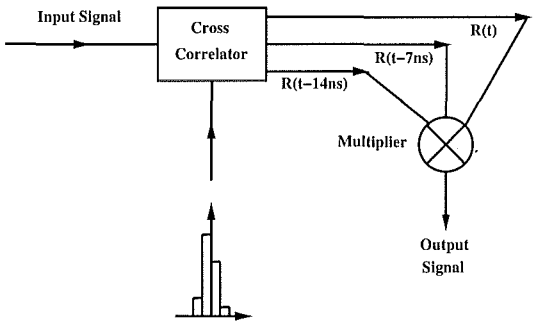


Fig. 3.18: Schematic of the correlation process used to improve the appearance of the lunar ranging data. The three channels that emerge from the cross correlator relate to the three main peaks observed in the raw data and are separated by 7 ns.

3.6 Parametric Models

As mentioned in several sections of this chapter the estimation of the spectrum of a signal is a central issue for the ring laser project. It has been shown that Fourier techniques have limitations in resolution whereas waveform fitting routines require explicit prior knowledge of the signals functional form and are so slow as to be prohibitive; the analysis of one data set (97.206 TST1-52) using simplex routines is estimated to require 80 days of processing on a Pentium Pro 200 dual processor machine using one technique. Clearly a fast method of estimating the power spectrum of our output is required. To this end parametric models are investigated here. The reason for choosing parametric models to describe the ring laser output waveform centres around the second order autoregressive model. This particular model describes a single spectral line of frequency f and width Δf which approximates a Lorentzian in shape when $\Delta f/f \ll 1$. Such a spectrum is precisely that which is predicted by quantum theory (see Chapter 4). The prior information provided by quantum theory is therefore concisely expressed in terms of this second order autoregressive model.

Parametric models describe stochastic processes. The most important model for this purpose is the second order autoregressive model, AR(2), which describes pseudo-sinusoidal oscillations. The AR(2) model provides a two-parameter description of an oscillating signal. These parameters may be mapped onto the two physically meaningful quantities of spectral line centre and spectral line width for the signal. It is desirable to have a measure of these quantities but in the presence of frequency drift it has been difficult to estimate them using the low resolution of the Fourier transform especially on the short data sets for which frequency drift is negligible. The AR(2) model requires only very short segments of data to provide an accurate estimate for both line centre and line width. For example when considering a band-passed Sagnac frequency from C-II this model provides an uncertainty in the estimation of the Sagnac frequency of ± 0.1 mHz for a segment of data of only 1 s. The other meaningful parameter is the linewidth estimate which is entirely inaccessible to the Fourier approach on such time scales.' The result of using such short data segments is that there are many estimates available. Information about their distribution can be formed and the parameters that describe this distribution are themselves information about the original signal. The time evolution of this distribution gives us an insight into the time evolution of the linewidth and hence to the underlying processes that give the spectral line width (most notably quantum noise).

The uses of parametric models in the description of stationary random processes [96, Chap 3] are considered here. An excellent text that covers this material is that due to Naidu [78]. A stationary random process is one which has statistical features that do not evolve with time, otherwise the process is called evolutionary. Parametric processes come in discrete and continuous parameter forms, both of which have very similar features. In their essence the discrete models are formed on a discrete time basis and are governed by difference equa-

tions, whereas the continuous parameter models are formed on a continuous basis and are thus described by differential equations. Although continuous parameter forms are mentioned they are not particularly applicable to the ring laser as they are not readily implemented on a computer.

In this chapter both the implementation and theory of parametric models are discussed. In order to get an understanding of the theory it is necessary to consider the equivalent techniques of parametric model estimation and the Maximum Entropy Method. The maximum entropy idea is a useful way of seeing how such high spectral resolution is achieved while the model fitting is far more easily implemented. An excellent review of the use of Maximum Entropy and Autoregressive methods is given by Ulrych and Bishop [117] in the very useful I.E.E.E. collection of papers edited by Childers [27] and its sequel by Kesler [65].

3.6.1 The Maximum Entropy Method

The autocorrelation function is central to the Maximum Entropy Method. Suppose a signal has a spectral component in it. An estimate of the spectrum may be provided by directly finding the Fourier transform. Alternatively, the indirect method may be used which evaluates the autocorrelation function and then makes use of the Wiener-Khintchine theorem to relate the autocorrelation to the power spectrum. It is common in signal processing to take this intermediate step because many signal enhancement techniques manipulate the autocorrelation function before forming the power spectrum. For a signal with a strong spectral component the autocorrelation function oscillates as the signal falls in and out of phase with itself for different time lags. The autocorrelation function is limited though by the length, N, of the data set used. The longest possible time lag is equal to the length of the data set. This restricts the resolution of the power spectrum. In effect the autocorrelation function is taken to be zero for lags longer than N; this is clearly not true. The autocorrelation coefficients need to be extended by some method. The use of maximum entropy to solve this problem was proposed by Burg [25]. The idea centres around the Jaynes principle [62] which is quoted from Ulrych and Bishop [117]

> The prior probability assignment that describes the available information but is maximally noncommittal with regards to the unavailable information is the one with maximum entropy.

Thus the autocovariance coefficients may be extended to time lags longer than the data set. This is done by considering the the first $N+1$ coefficients. The entropy of the signal is given by Eq. (3.31) but with $C(N+1)$ replacing $C(N)$. The only unknown is $\hat{\rho}(N+1)$ which is varied so as to maximise the entropy H. The process continues indefinitely to the required time lag by iterating this procedure. The power spectral density estimate is then calculated using the new enlarged sequence of estimates for the autocovariance and has a great potential

for improved spectral resolution. The above process needs fuller explanation and to do this a brief summary of the relevant information theory is needed.

Information Theory

Information theory deals with the information content of signals. It is common to think of information as conveying meaning but in this field this is not the case. Information is conveyed by a signal which can take many different values.

Consider a finite sequence of characters; this sequence may have W different realizations. A particular realization isolates one message from all the others; information has been received. Information is thought of as being related to the isolation of one message from all the possible messages. Further, a message of twice the length ought to carry twice the information while having W^2 different realizations. This indicates that the definition of information is related to the logarithm of W. In general allowance must be made for the fact that different realizations are not necessarily of equal likelihood. The general definition of information is

$$H = -k \sum_{i=1}^{N} p_i \log p_i \tag{3.27}$$

where k is constant equal to one if the base of the logarithm is 2. This is the only definition of H that satisfies the reasonable restrictions which are naturally placed on a measure of information. This idea is principally due to Shannon and is found in Shannon and Weaver [103]. This definition is equivalent to the definition of entropy and H is called the entropy of the signal.

A comparison with the methods of statistical mechanics can be found by considering the entropy[3]

$$S = -k \sum_{i,N} p_{iN} \ln p_{iN} \ . \tag{3.28}$$

In particular consider the micro-canonical ensemble description of an ideal gas where the number of particles N and the internal energy of the gas are known. There is no reason for the system to be in one viable location in phase space rather than another; all the states are equally likely so that $p_{iN} = p_i = 1/W$. Hence

$$S = -k \ln \frac{1}{W} \sum_{i,N} p_{iN} = k \ln W \ . \tag{3.29}$$

Here the entropy is related to $\ln W$ as suggested earlier but it is required that all states are equally likely, a situation which corresponds with the maximum entropy or maximum disorder for the system or in the case of a signal, white noise.

[3] Notice that in statistical mechanics the natural logarithm is used and the constant is Boltzmann's constant which gives the proper units for this entropy measure.

The Extension of the Autocorrelation Function

The extension of the autocorrelation coefficients beyond time lags equal to the length of the data set requires the maximisation of the signal entropy with respect to those coefficients. Thus the signal entropy needs to be written in terms of the autocorrelation coefficients. This is done in a series of steps that follow Ulrych and Bishop [117]. The entropy of a signal may be re-expressed in terms of the the power spectrum according to Smylie et. al. [106]

$$H = \frac{1}{4f_N} \int_{-f_N}^{f_N} \log S(f) df \ . \tag{3.30}$$

and rewriting in terms of the autocovariance estimates:

$$H = \tfrac{1}{2} \log \left(\det C(N) \right) \tag{3.31}$$

where $C(N)$ is the Toeplitz autocovariance matrix

$$C(N) = \begin{bmatrix} \hat{\rho}(0) & \hat{\rho}(1) & \cdots & \hat{\rho}(N) \\ \hat{\rho}(1) & \hat{\rho}(0) & \cdots & \hat{\rho}(N-1) \\ \vdots & \vdots & & \vdots \\ \hat{\rho}(N) & \hat{\rho}(N-1) & \cdots & \hat{\rho}(0) \end{bmatrix} . \tag{3.32}$$

There exist several different methods of evaluating the first few autocovariance estimates, starting with

$$\hat{\rho}(0) = \frac{1}{N} \sum_{i=1}^{N} {x_t}^2 \tag{3.33}$$

the rest are found up to $\hat{\rho}(N)$. One of these methods, the Burg method, is considered later in §3.6.4. Prior to the development of the current method the autocovariance coefficients would be used to estimate the power spectrum either by directly Fourier transforming or by filtering the coefficients and then transforming.

3.6.2 Discrete Parameter Models

In their essence parametric models describe random processes where there is some dependence between neighbouring samples. Consider, for example, the current exiting a photodiode, upon detection the current rises sharply and then decays exponentially with a characteristic time constant. Now the arrival of pulses is usually a Poisson distributed process and is random. If the current is sampled at such a rate that another sample is taken before the current has decayed effectively to zero then successive samples may be expected to obey a relation

$$X_t + aX_{t-1} = \epsilon_t \tag{3.34}$$

where the X_t are the samples, a is a parameter and ϵ_t is a purely random process; usually the random process is N(0,1) which is a Gaussian distribution with zero mean and unit variance. The process described above is known as a first order autoregressive process AR(1) or sometimes as a Linear Markov process.

There are several well known parametric models, these are autoregressive AR(k) models

$$X_t + a_1 X_{t-1} + \cdots + a_k X_{t-k} = \epsilon_t \tag{3.35}$$

moving average MA(l) models

$$X_t = \epsilon_t + b_1 \epsilon_{t-1} + \cdots + b_l \epsilon_{t-l} \tag{3.36}$$

and combined autoregressive moving average models ARMA(k, l)

$$X_t + a_1 X_{t-1} + \cdots + a_k X_{t-k} = \epsilon_t + b_1 \epsilon_{t-1} + \cdots + b_l \epsilon_{t-l} \ . \tag{3.37}$$

In general when analysing the ring output only short data segments are considered during which drift is negligible. Hence the segments may be assumed to be stationary stochastic processes and only autoregressive models need be considered. It is conceivable that the nature of the carrier drift could be modelled using an ARMA model; this has not been considered here.

3.6.3 Continuous Parameter Models

The discrete parameter models have a natural extension into the continuous time domain. The extension gives rise to stochastic differential equations which form a large field of study which is not entered into in any depth here. However, an overview is included with reference only to the second order model. Referring to Eq. (3.48) the difference equations

$$\dot{X}(t) = X_t - X_{t-1} \tag{3.38}$$

$$\ddot{X}(t) = X_t - 2X_{t-1} + X_{t-2} \tag{3.39}$$

are used where the sample points are taken to be separated by unit time. Thus the equation that defines the AR(2) process can be transformed to the second order stochastic differential equation

$$\ddot{X}(t) + \alpha_1 \dot{X}(t) + \alpha_2 X(t) = \epsilon(t) \tag{3.40}$$

which is precisely the equation that defines the continuous parameter AR(2) model. This equation can be immediately identified as a damped harmonic oscillator driven by an entirely random driving force. A resonance is found which depends on the values of the coefficients of the model. Further comments on the solution of stochastic differential equations that are relevant to the ring laser are found in §4.2.2.

3.6.4 Burg Estimates of Autoregressive Model Coefficients

The method of maximum entropy has been shown by Van den Bos [119] to be mathematically equivalent to least squares fitting of autoregressive models to the data. The method of evaluating the Burg spectral estimate involves fitting autoregressive models to the data and determining the spectrum from the model coefficients. These Burg estimates are made in a fashion that is consistent with the maximum entropy method.

When a data sequence is to be described in terms of a discrete parameter model there is one question that has to be answered: What order model should be used? The simplest method is simply to look at the autocovariance function and compare it with some library of known functions. Alternatively a series of models can be tested to see if one gives a markedly better fit than the others. Here a selection of successively higher order autoregressive models are analysed and the quality of fit found, from this the appropriate order of the model is estimated. In the scheme used here a prediction error for the first order model is evaluated and an iterative process uses this value to find the prediction error for all models up to a length equal to the length of the data set. The appropriate model is chosen from this prediction error. The prediction error used is that due to Akaike [2, 3] and is called the Final Prediction Error (FPE) defined by

$$\mathrm{FPE} = E[(x_t - (\hat{x}_t))^2] \; . \tag{3.41}$$

In the Burg scheme the FPE for a model of order m and a data set of length N which has the mean subtracted is given by

$$(\mathrm{FPE})_m = \frac{N + (m+1)}{N - (m+1)} P_m \; . \tag{3.42}$$

Here P_m is the minimum sum of residual squares which according to the Burg scheme is found by passing the prediction filter, the coefficients of the autoregressive model, forward and backwards over the data according to

$$\begin{aligned} P_{m+1} \;=\; & \frac{1}{2(N-m)} \sum_{m+1}^{N} \{(x_t + \hat{a}_{m1} x_{t-1} + \cdots + \hat{a}_{mm} x_{t-m})^2 \\ & + (x_{t-m} + \hat{a}_{m1} x_{t-m+1} + \cdots + \hat{a}_{mm} x_t)^2\} \end{aligned} \tag{3.43}$$

where the $\hat{a}_{ij}$ represent the j^{th} component of the prediction filter for the model of order i. The parameters are determined by minimising the sum of residual squares. It is straightforward to show that the first order Burg estimate is given by

$$\hat{a}_{11} = -\frac{\sum_{t=2}^{N} x_t x_{t-1}}{{x_1}^2/2 + \sum_{t=2}^{N-1} {x_t}^2 + {x_N}^2/2} \; . \tag{3.44}$$

The value for the FPE is calculated and efficient coding allows the incumbent value $\hat{a}_{11}$ to be used in the evaluation of the next filter values $\hat{a}_{21}$ and $\hat{a}_{22}$. This process is iterated until the filters and FPEs are found for all orders up to the length of the data set. The power spectrum estimate is then given in terms of the filter coefficients of an m^{th} order AR process according to

$$S_m(f) \propto \frac{\sigma_\epsilon^2}{\left|1 + \sum_{j=1}^{m} \hat{a}_{mj} \exp(-i2\pi f j \Delta t)\right|^2} \tag{3.45}$$

where Δt is the uniform sampling rate.

An Illustrative Example

In order to illustrate the ideas discussed here a fourth order AR process is analysed which is taken from Ulrych and Bishop [117, p.193] and has coefficients

$$\{2.7607, -3.8106, 2.6535, -0.9238\} \,. \tag{3.46}$$

A data set that is generated by this mechanism is shown in Fig. 3.19. It can be seen that this process is not purely harmonic as the periods of the various cycles are clearly not equal. There is also an apparent amplitude envelope which may indicate the beating of two close frequencies. These observations are purely qual-

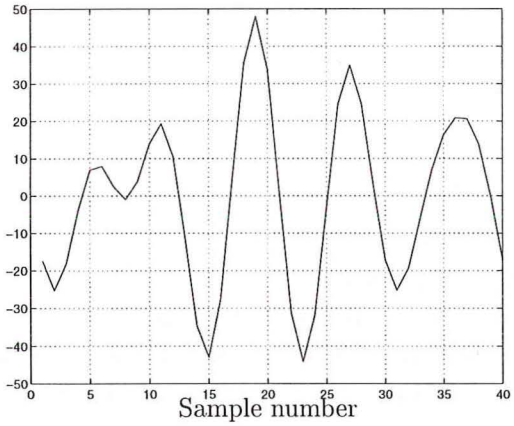


Fig. 3.19: A typical realization of the fourth order process which comprises forty data points.

itative but the analysis outlined above makes them quantitative and ultimately leads to a spectral estimate. The order of the process which best describes the data is determined. Fig. 3.20 shows how Akaike's FPE varies with the order of the model being used. Notice that for orders above four there is no real improvement but the error is also no worse. It is natural to choose the model of lowest order that gives the best fit. In this case it is clear which order to assume. As the order of the model approaches the length of the data set the FPE becomes unreliable; a model with order much less than the length of the data set is desired

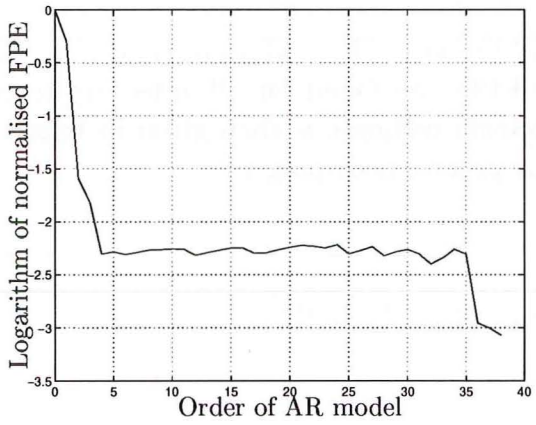


Fig. 3.20: This figure shows final prediction error according to the Akaike scheme for a data set of 40 samples that is known to be generated by a fourth order autoregressive process.

anyway. Having settled on the fourth order model the coefficients can be inserted into Eq. (3.45) and the spectral estimate found. This is illustrated in Fig. 3.21. The spectral estimate is directly contrasted with the known spectrum deduced

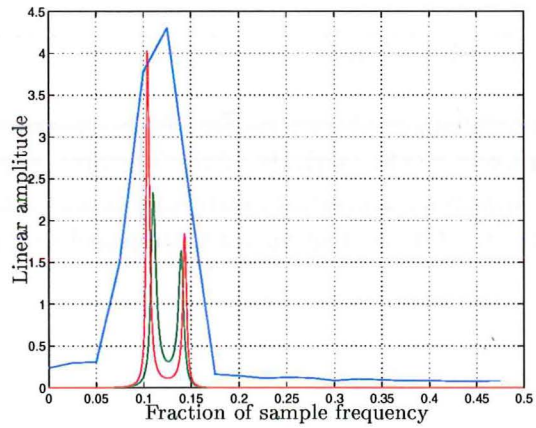


Fig. 3.21: The green spectrum is the true spectrum which is deduced from the coefficients of the AR process, the red line shows the fourth order AR fit to the data and the blue line shows the direct Fourier transform of the data.

from the coefficients given in Eq. (3.46) and the discrete Fourier transform. The resolution of the Burg method is a great improvement over the Fourier method and is close to the known spectrum. However, because the data set is generated by a random process there are many such realizations that do not achieve this same degree of resolution and sometimes fail to see two spectral lines. The Burg estimate is never worse than the Fourier method for realizations of this length.

Later, when the ring laser signal is analysed, a reliable estimate of the AR parameters will be required. The fourth order example is examined further as in this case the exact values of the parameters are known. Many different realisations of the fourth order process are analysed and a distribution found for each of the four parameters. One such distribution is shown in Fig. 3.22. To retrieve an accurate parameter value from this data requires a fit of some functional form. It is found that a function of the form

$$f(x) = \kappa \frac{e^{-\mu}\mu^{\lambda x-\rho}}{\Gamma(\lambda x + 1 - \rho)} \tag{3.47}$$

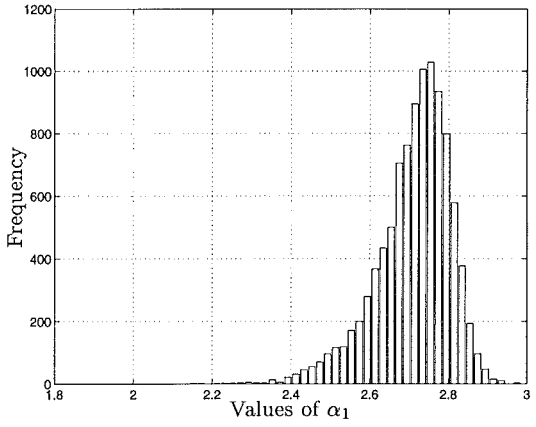


Fig. 3.22: This histogram shows the predicted values of the first coefficient of the fourth order model which results from 10,000 realizations of the process (±1%). The actual value is known to be 2.7607 whereas the mean of all the samples is 2.7103 and the median is 2.7257. However the maximum of the distribution is close to the expected value.

fits the data well. This is a Poisson distribution with mean μ, displacement ρ and scale factor κ. The λ parameter is either ± 1 and indicates whether the Poisson distribution is reflected about the x-axis ($\lambda = -1$ in the example above). The estimate of the α parameter is given by the location of the stationary point of this distribution. The parameters κ, λ, ρ and μ are varied to give a best fit to the histogram of the data and the value of α found. This method is entirely empirical and there is no theoretical reason for selecting the functional form given in Eq. (3.47). However, as is shown in the table below, the α estimates do tend to the known values for large numbers of realizations of the process. When this mechanism for determining the α values is applied to the fourth order model estimate the predicted parameter values are as follows:

known value	estimate from 10^4 segments		estimate from 10^5 segments	
2.7607	2.7469	δ =0.0138	2.7660	δ =-0.0053
-3.8106	-3.7905	δ =-0.0201	-3.8103	δ =-0.0004
2.6535	2.6321	δ =0.0214	2.6497	δ =0.0038
-0.9238	-0.9254	δ =0.0016	-0.9252	δ =0.0014

3.6.5 The AR(2) Model

For the purposes of the ring group the most important model is the AR(2) model which is analysed in greater depth here. This work follows Ulrych and Bishop [117]. The model is given by

$$x_t + a_1 x_{t-1} + a_2 x_{t-2} = \epsilon_t \ . \tag{3.48}$$

Quite often a process may require some time for transient features to die away and thereafter the process is stationary. This is known as asymptotic stationarity. In the AR(2) model this is satisfied if the values for a_1 and a_2 lie within the triangle defined by $a_2 \leq 1$, $a_1 + a_2 \geq -1$ and $a_1 - a_2 \geq 1$. The different types of behaviour are classified by means of the characteristic equation

$$z^2 + a_1 z + a_2 = 0 \tag{3.49}$$

which has roots μ_1 and μ_2. The autocorrelation function is given by the expression

$$R(r) = \frac{(1-{\mu_2}^2){\mu_1}^{r+1} - (1-{\mu_1}^2){\mu_2}^{r+1}}{(\mu_1-\mu_2)(1+\mu_1\mu_2)} \tag{3.50}$$

where the index $r \geq 0$ and by symmetry of the autocorrelation function $R(r) = R(-r)$. If the roots μ_1 and μ_2 are both real and positive then the autocorrelation function has a bell shape; if the roots are real and negative then the autocorrelation function has the same shape but alternates in sign at each successive value of r. When the roots are complex ($-\alpha_2 > {\alpha_1}^2/4$) and the autocorrelation function may be rewritten as

$$R(r) = {a_2}^{r/2}\left(\frac{\sin(r\theta+\psi)}{\sin\psi}\right) \tag{3.51}$$

with

$$\tan\psi = \left(\frac{1+a_2}{1-a_2}\right)\tan\theta \quad \text{and} \quad \cos\theta = \frac{-a_1}{2\sqrt{a_2}}\,. \tag{3.52}$$

which has the appearance illustrated in Fig. 3.23. The presence of periodic os-

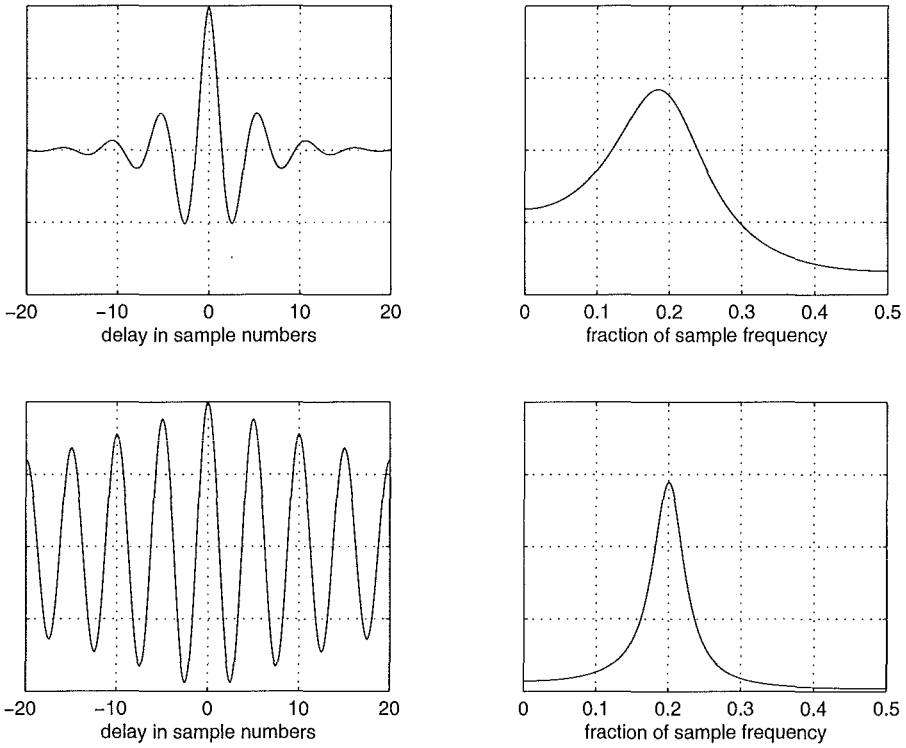


Fig. 3.23: The autocorrelation functions (left) and predicted spectra (right) for complex roots arising from the AR(2) parameters (top two plots) a_1=-0.6 and a_2=0.6 and (bottom two plots) a_1=-0.6 and a_2=0.95. Notice that continued correlation for longer delays equates to narrower linewidths.

cillations in the autocorrelation function indicates that there is a periodicity in

the time domain data. However, the attenuation of the autocorrelation tells us that the correlation disappears as the lag grows so that there is a characteristic coherence time associated with the data; alternatively the process is said to be *pseudo-periodic*. Such behaviour is incoherent and gives rise to a broad spectral feature with width associated with the degree of coherence. This is illustrated in Fig. 3.23 where the theoretical power spectrum shows the lineshape resulting from the autocorrelation function. The power spectrum of the AR(2) process is explicitly

$$S(f) = \frac{1}{2\pi} \frac{2\sigma_\epsilon^2}{1 + a_1^2 + a_2^2 + 2a_1(a_2+1)\cos(2\pi f) + 2a_2\cos(4\pi f)} \tag{3.53}$$

with the peak occurring at

$$f_0 = \frac{1}{2\pi}\cos^{-1}\left(\frac{-a_1(a_2+1)}{4a_2}\right) \quad \text{provided} \quad a_2 > 0 \quad \text{and} \quad \left|\frac{a_1(a_2+1)}{4a_2}\right| < 1 \tag{3.54}$$

and which has full width at half-maximum in hertz

$$\Gamma_{\mathrm{AR2}} = \frac{1}{2\pi}\left(\cos^{-1}(\cos(2\pi f_0) - \zeta) - \cos^{-1}(\cos(2\pi f_0) + \zeta)\right) \tag{3.55}$$

where

$$\zeta = \frac{\sqrt{4a_2 - a_1^2}(1 - a_2)}{4a_2} \tag{3.56}$$

This linewidth is expressed in hertz where the sampling frequency is normalised to 1 Hz. In general f_0 and Γ_{AR2} will need to be scaled by the sampling frequency.

The numerical determination of the coefficients is achieved by minimising the sum of square residues which arise from the Burg scheme, Eq. (3.43). The estimates for the coefficients can be found from

$$\begin{aligned} 2\langle 0,2\rangle &= -a_1\left(\langle 1,2\rangle + \langle 1,0\rangle - 2\frac{\langle 1,1\rangle(\langle 0,0\rangle + \langle 2,2\rangle)}{\langle 2,1\rangle + \langle 0,1\rangle}\right) + \langle 0,0\rangle + \langle 2,2\rangle \\ 1 &= -2a_1\frac{\langle 1,1\rangle}{\langle 2,1\rangle + \langle 0,1\rangle} - a_2 \\ \text{and} \quad \langle \alpha,\beta\rangle &= \sum_3^N X_{t-\alpha}X_{t-\beta} \end{aligned}$$

The AR(2) model provides a description of a pseudo-sinusoidal signal with a spectral lineshape that tends towards a Lorentzian for narrow spectral lines. Thus the model provides an appropriate description of the ring laser output. The innovation made in this thesis is to consider Γ_{AR2} as serious estimate of the quantum noise induced linewidth. The estimate requires scaling, the magnitude of this scaling is determined by the characterisation procedure described in the next section.

3.6.6 Characterisation of the AR(2) Method

In order for the AR(2) model to provide an appropriate description of the ring laser output it is necessary that the only significant spectral feature is a single spectral line. Unfortunately real data sets contain many spurious signals as well as the harmonics of the Sagnac frequency. These contaminating features are easily removed using a digital band pass filter, width B, in post-processing; but the validity of using the AR(2) model in analysing the resulting filtered data is not assured.

A correction factor for the linewidth estimates from a given ring laser data set is required. This correction factor is determined by producing artificial data sequences which match the ring laser output as accurately as possible. These synthesised data sequences contain a spectral line at the ring laser frequency with a precisely known spectral linewidth. A broadband white Gaussian noise sequence is also added to the signal so that it matches the power spectrum of the ring output. The resulting artificial data is analysed using the AR(2) model and the linewidth estimates compared with the known linewidth. In this manner a scaling factor is found which corrects the AR(2) linewidth estimates.

Synthesised waveforms are created according to the formula:

$$\sin(2\pi f_0 t_i + \phi(t_i)) \quad \text{where} \quad i = 1, 2, \ldots$$

where the time steps are evenly spaced with interval τ. White frequency noise is introduced into the signal in the form of a random walk in phase $\phi(t_i)$ which is defined as

$$\phi(t_i) = \sum_{k=1}^{i} \epsilon_k \quad \text{and } \epsilon_k \text{ are distributed according to} \quad P(\epsilon) = \sqrt{\frac{k}{\tau\pi}} e^{-k\epsilon^2/\tau} .$$

The probability distribution has been scaled so that the rate of phase diffusion is independent of τ. Now a relationship is required between the rate of phase diffusion, defined by k, and the full-width at half maximum linewidth in hertz, Γ_{Hz}, which this phase diffusion induces. Given a measurement time T of N steps, so that $T = N\tau$, the probability distribution of the total phase diffusion $\phi_N = \epsilon_1 + \epsilon_2 + \cdots + \epsilon_N$ is given by (see Appendix D)

$$P(\phi_N) = \sqrt{\frac{k}{N\tau\pi}} e^{-k\phi_N^2/N\tau} \tag{3.57}$$

and the mean and variance of the distribution are $\langle\phi_N\rangle = 0$ and $\langle\phi_N^2\rangle = T/2k$. The root mean square frequency fluctuations are defined as

$$\Delta\omega_{\text{rms}} = \frac{\Delta\phi_{\text{rms}}}{T} = \frac{\sqrt{\langle(\phi_N - \langle\phi_N\rangle)^2\rangle}}{T} = \frac{1}{\sqrt{2kT}}$$

which is compared with the standard result $T\Delta\omega_{\text{rms}}^2 = \Gamma_{\text{rad}}$ where Γ_{rad} is the full width at half maximum linewidth in radian measure (see for example Statz

et al. [107]). Clearly then the linewidth and the phase diffusion are related by $\Gamma_{\text{rad}} = 1/2k$. When a real implementation is made the phase diffusion is provided by a term $\Delta\phi N_i$ where N_i are normally distributed with centre zero and unit variance. The root mean square phase diffusion in unit time is $(\Delta\phi)^2/\tau$. Finally the relation between the magnitude of the random numbers used in the phase diffusion, $\Delta\phi$, and induced linewidth is $\Gamma_{\text{Hz}} = (\Delta\phi)^2/2\pi$. Thus artificial data sequences are generated for which the linewidth and line centre are known exactly; the ability of the AR(2) model to estimate these parameters may then be assessed.

An additional noise sequence is added to the artificial signal which is intended to represent background white noise in the detection system. This noise is taken to be zero mean and variance β^2 so that the sequence $\beta N(i)$ is added to the signal. The value of β is determined by matching the power spectra of the synthesised and real data sequences. The amplitude of the noise on the Sagnac signal is never greater than 10% of the peak to peak amplitude of the Earth induced oscillation.

4. Noise in Ring Lasers

In this chapter the effects of noise on the ring laser output Sagnac frequency are discussed. A variety of different noise types are considered from instrument noise to quantum noise which may potentially create uncertainty in the Sagnac frequency. The magnitude of this uncertainty is explored along with the theoretical minimum that quantum mechanics places on it. An Allan variance analysis of the output frequency from the ring lasers suggests that for short integration times they operate at the quantum limit but for longer integration times environmentally induced frequency drift dominates. The white frequency noise induced linewidths are not large enough to explain the departure of the measured Sagnac frequency from the theoretically expected frequency.

4.1 Sources of Instrument Noise

In this section only those noise sources that are relevant to data collection from the ring lasers are discussed. The records with the lowest signal noise from C-I (for example the data set that I took in December 1995 T95DE18b.sc1) give an audio Sagnac tone which rises over 60 dBV above a white noise floor when displayed on the SRS 770 audio spectrum analyser with a resolution bandwidth of 1 Hz. In order to determine precisely the rms noise voltage on this signal a sinusoid and harmonics are fitted to the output signal and subtracted from it leaving only the white noise background. The residual has an rms fluctuation of 130 A/D voltage levels with the signal having a peak to peak amplitude of 4840 A/D voltage levels. The A/D converter is 14 bit with a ± 8 V range. Thus the signal has an amplitude of $4.7\,V_{p\text{-}p}$ with a white noise content of $127\,mV_{rms}$. In practice the variation of operating parameters from one data collection sequence to the next is such that only the rough estimate can be made that the signal noise is of the order of several hundred mV. In this section an account of instrument noise is made in order to allocate the proper fraction of this signal noise to the detection system.

A noise spectrum which has constant average power up to a particular frequency is often referred to as band-limited white noise. This arises from an optical analogue where white corresponds to a flat spectral density across the visible spectrum. The analogy is extended to other types of noise. Noise which is passed through a filter with non-constant transfer function is called coloured noise while noise with a $1/f$ spectrum is referred to as pink noise (or simply $1/f$-noise) due to its

stronger spectral components at the red end of the spectrum.

A schematic of the arrangement of the detection system is given in Fig. 4.1. All

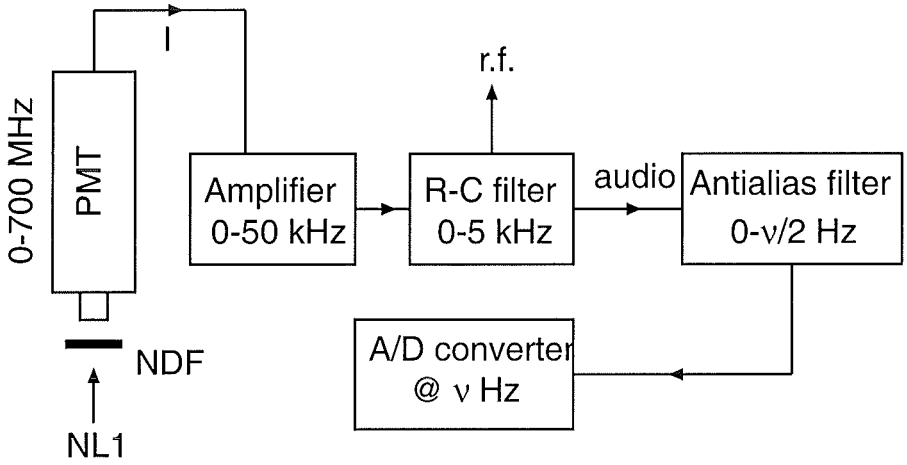


Fig. 4.1: A schematic of the detection scheme used. I introduced the antialias filter in March 1997 in order to minimise the noise on the final signal. The relevant bandwidths at various stages of the process are included.

the noise terms added at the various stages are white up to the effective bandwidth at that stage. The first noise term that is added in the detection process is shot noise in the photomultiplier and extends up to 700 MHz cut-off frequency. This noise current is amplified by the transimpedance amplifier which also adds its characteristic noise and has a passband of 0-50 kHz. Following the amplifier is a separate R-C filter with 0-5 kHz passband. The recent addition of the antialias filter brings the passband down to half the sampling frequency and prevents high frequency noise being aliased down into the Nyquist window, which unnecessarily adds to the background noise. A broad discussion of noise sources and low noise techniques is given by Horowitz and Hill [59, Chapter 7].

4.1.1 Shot Noise

Shot noise refers to a stochastic process that is identified with distinct events. Typical devices that give rise to shot noise include photon and electron detectors and various vacuum valves or, more generally, any particle detector. The events in these devices usually initiate the flow of a quantity of charge and hence the shot noise appears in the output current of the device. The events occur at random times τ_i and the number of events $N(t)$ that occur in a time interval $(0, t)$ is usually given by the Poisson distribution $p(N(t) = i) = e^{-\mu}\mu^i/i!$ where the mean and variance are both μ. In practice the detector response cannot be instantaneous and may have a time lag and perhaps decay characteristics; an event at time zero gives rise to an output $g(t)$. When several such responses are superimposed they are assumed to combine linearly so that the output current is

written

$$I(t) = \sum_{i=1}^{N(t)} g(t - \tau_i) \ . \tag{4.1}$$

For sufficiently high count rates, as with C-I, the individual counts are not measured, rather, the macroscopic anode current is the working variable. The spectral density is that of a white noise spectrum due to the stochastic nature of the arrival times of the photons. This white spectrum tapers off at a frequency defined by the workings of the particular detector (our Hamamatsu Photomultiplier tubes begin to attenuate at $\sim 700\,\mathrm{MHz}$ but can still detect signals at $\sim 1\,\mathrm{GHz}$). Other sources of shot noise in electronics such as transistor and FET noise are not considered here as they are only of limited relevance to the project but significant discussion of these devices can be found in Connor [30] along with several technical references.

4.1.2 Thermal Noise and Noise Figure

Thermal noise relates to the noise that arises in, for example, a resistor due to thermal motion of the charge carriers and is often referred to as Johnson noise. The voltage fluctuations across a resistor are given by

$$V_{\mathrm{rms}} = \sqrt{4k_BTBR} \quad \mathrm{V} \quad \text{or} \quad V_{\mathrm{rms}} = \sqrt{4k_BTR} \quad \mathrm{V}/\sqrt{\mathrm{Hz}} \tag{4.2}$$

where k_B is Boltzmann's constant, T is the absolute temperature, B is the bandwidth of the detector and R is the value of the resistance. The spectral density of this white noise voltage is given by $S_V(f) = 2k_BTR$ and is constant up to $\sim$13 MHz after which it tapers off.

4.1.3 A/D Conversion Noise

The last source of noise considered is A/D conversion noise. This occurs as a result of the discrete nature of analogue to digital (A/D) conversion. The A/D converter equates the input voltage with one of a discrete set of voltage bands each with width ΔV which cover the interval -8 V to 8 V and hence an ideal digital representation of the input signal is not possible. For a low amplitude signal the discrete levels will be quite apparent. This effect imposes a limit on the best attainable signal to noise ratio which is given by

$$-20\log_{10}(2^{2p} - 1) \quad (\mathrm{dB}) \simeq -40p\log_{10}(2) \quad (\mathrm{dB}) \tag{4.3}$$

for a p-bit A/D converter. [30, p.89 et. seq.] This limit is only achieved when the signal has a peak to peak amplitude equal to the full dynamic range of the A/D converter. There exists a technique [59] that may be used to improve this signal to noise ratio which allows a voltage resolution better than the discrete

voltage levels, ΔV. The technique involves the deliberate addition of noise to the signal. A small noise voltage is added to the signal and the signal is oversampled and averaged. The average of many readings reduces the noise and gives a better resolution of the input voltage. It is found that the process works best for additive noise of amplitude $\pm\frac{1}{2}\Delta V$. To illustrate this idea consider a sinusoidal voltage which has insufficient amplitude to span the quantisation levels, the output would at best be a series of measurements which alternated between two levels but may not change at all. When the noise is added any sample will have at least two possible outcomes so that when many samples are taken (at a rate very rapid relative to the signal frequency and for a duration much less than a period) and averaged an accurate estimate of the signal value at that time may be found. This estimate may have variance much smaller than ΔV.

4.1.4 Detection System Noise

Working through the detection system the magnitudes of the various noise contributions are found. In this way an idea of the total noise contribution is found along with the identification of the particular elements that are responsible for them. Recall that the output noise observed on the data set T95DE18b was of the order of a few hundred mV. The noise characteristics of the PMT, the high tension (HT) power supply and the amplifier are taken. The amplifier is connected to the SRS 770 network analyser and its output spectrum observed from 0-100 kHz. With the HT set to 0 V the white noise floor is $15.6\,\mu\mathrm{V}/\sqrt{\mathrm{Hz}}$. The amplifier has a following R-C filter with the 3 dB point at 50 kHz and hence $3.5\,\mathrm{mV_{rms}}$ is present at the output of the amplifier. Next the HT was set to the typical operating value of 800 V and the input laser beam attenuated until a mid range (few volts) dc output voltage was observed from the amplifier; this emulates the operating conditions of C-I. The photomultiplier tube used was the R3896 from HAMAMATSU and has a typical dark current of approximately 10 nA. The working range of this PMT has a maximum current of 0.1 mA when operated at full scale. With the input beam occulted the spectrum has a white noise floor at $142\,\mu\mathrm{V}/\sqrt{\mathrm{Hz}}$ with the same 50 kHz passband. The output noise that passes through the network is thus $32\,\mathrm{mV_{rms}}$. With the beam allowed to pass through to the PMT the noise level rose to $1.5\,\mathrm{mV}/\sqrt{\mathrm{Hz}}$ equating to $350\,\mathrm{mV_{rms}}$ in the amplifier output. The conclusion is thus made that in typical operation of the order of 1% of the noise at the amplifier output is due to the amplifier itself. Approximately 10% of this noise arises from the PMT when powered at 800 V and the remaining 90% is thus due to fluctuations in the input beam irradiance.

For the data set under consideration (T95DE18b.sc1) the antialias filter was not used and hence a final passband of ~5 kHz was present due to the R-C filter. The sampling frequency was 500 Hz so that the Nyquist window was 250 Hz wide. In the absence of the antialias filter the noise passed through the 0-5 kHz window is repeatedly aliased, or folded, into the Nyquist window. This increases the noise floor by $\sqrt{n}$ where n is number of times the noise is folded into the Nyquist

window; in this case $n = 20$. Thus a noise voltage of $\sim$470 mV_{rms} is expected according to the above measurements. The introduced rms noise may be roughly stated as a few hundred mV.

The antialias filter is also analysed on a network analyser and performs close to specification. The white noise floor is below -100 dBV with a resolution bandwidth of 1 Hz and hence accounts for $<10\,\mu V/\sqrt{Hz}$ of rms noise. The filter uses an internal clock to determine the filter cutoff frequency which oscillates at 100 times the bandpass frequency. A small amount of this clock signal leaks into the output and can be seen at -90 dBV and thus provides a 30 μV tone at the clock frequency. The noise contributions from the antialias filter are negligible relative to the noise seen after the amplifier stage.

Quantisation noise depends on how much of the dynamic range of the A/D converter is used. Typically a large fraction (at least a quarter) of this range is used; the full scale is not utilised so that the amplitude drift may be contained without clipping. Using the stated formula (§4.1.3) for the STROBES 14 bit A/D converter the introduced noise floor is $\sim$-170 dB below the carrier and is also negligible.

The measured output rms noise is measured at a few hundred mV. For our detection system this is divided between the PMT dark signal and the input intensity detection fluctuations in the ratio 1:9. The introduction of an antialias filter prevents noise being folded into the Nyquist window. With such a filter in place the 1.5 $mV/\sqrt{Hz}$ noise may be post processed through a narrow band digital filter centred on the carrier frequency (I use 16 Hz bandwidth) to give a $\pm$10 mV_{rms} noise signal on a Sagnac signal of 5 $V_{p\text{-}p}$. This represents an excellent signal to noise ratio which does not interfere appreciably with the AR(2) frequency determining method discussed in Chapter 3. This background noise does play a significant part in the linewidth estimation discussed in §4.4. Thus detection system noise, under most circumstances discussed here, is not an important factor in our measurements.

4.2 Quantum Noise

In a lasing cavity the gain medium gives rise to a background spontaneous emission which is a source of shot noise. In addition, at the lasing frequency there is necessarily spontaneous emission. There always exists some spontaneous emission at exactly the lasing frequency some of which will be in the direction of lasing but which will have random phase with respect to the lasing mode. Later (§4.2.2) it is seen that this can be described by a complex field strength of constant amplitude which in addition to its normal rotation at the lasing frequency has a phase noise component. This phase noise comprises a random walk in phase as the phase contribution to the mode at each event effects the mode permanently. Thus, if phase contributions occur at times τ_i the random walk in phase space may be

reasonably expressed as

$$\phi(t) = \sum_{i=1}^{n(t)} \Delta\phi g(t - \tau_i) \tag{4.4}$$

where $n(t)$ is the number of events up to time t, $g(t)$ is the response for a single event at time $t = 0$ and $\Delta\phi$ is a measure of the size of effect that one event has on the laser field strength. It is required that $g(t) = 0$ if $t < 0$ and $g(t) \to 1$ as $t \to \infty$; it usually suffices to let $g(t)$ be a unit Heaviside function. The field strength is finally given by $E(t) = E_0 e^{i\omega t + \phi(t)}$. It should be noted that white frequency noise is entirely equivalent to a random walk in phase space. This follows from the idea that a fluctuation of the frequency in one measurement gives rise to a phase contribution that is present in all the subsequent measurements. Hence white frequency fluctuations give white phase noise contributions that are cumulative. This is the characteristic of a random walk in one dimension (see Appendix D).

§4.3.5 suggests an experiment to determine the precise nature of the phase difference between the two counter-propagating beams in a ring laser gyro. In the absence of a ring laser specific phase theory the standard theory from homodyne methods is used. Phase uncertainty remains the source of the fundamental limits on the precision that may be attained by ring laser gyros. This section analyses these limits where it is assumed that phase uncertainty is described by the number-phase relation as discussed in §4.3.5. Two further developments of quantum noise are pursued in §4.2.2. These provide vital insights into experimental aspects that may obscure access to the quantum limit. Once these aspects are taken into consideration it is clear that all three analyses of the quantum limit agree. In §4.2.3 non-classical fields, that is not coherent, are discussed which have the property of potentially surpassing the above quantum limit. Practical realizations of the quantum limit are discussed in §4.4.3 and are contrasted with the performance of the local ring gyros.

4.2.1 Linewidth from the Uncertainty Relations

An initial derivation of linewidth is presented which starts with the uncertainty relation, $\Delta n \Delta\phi \geq 1/2$, as mentioned earlier; this is the approach taken by Dorschner et al. [41] and adopted later by Statz et al. [107]. When the equality holds the system is said to be in a minimum uncertainty state. There are many different combinations of number states that are minimum uncertainty states [58]. Here, only the coherent states, due to Glauber [51, 52], are considered as these properly describe the laser mode:

$$|\alpha\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_n \frac{\alpha^n}{\sqrt{n!}} |n\rangle$$

The uncertainties in the measurement of n may be characterised by root-mean-square deviations from the average according to

$$\Delta n = \langle (n - \langle n \rangle)^2 \rangle^{1/2} = \left(\langle n^2 \rangle - \langle n \rangle^2 \right)^{1/2} . \tag{4.5}$$

The following results hold for coherent states and are quoted from Loudon [72, pp.148]:

$$\begin{aligned} \langle n \rangle &= \langle \alpha | \, \hat{n} \, | \alpha \rangle = |\alpha|^2 \\ \langle n^2 \rangle &= \langle \alpha | \, \hat{n}^2 \, | \alpha \rangle = |\alpha|^4 + |\alpha|^2 . \end{aligned} \tag{4.6}$$

Collecting expressions it is found that $\Delta n = \langle n \rangle^{1/2}$ which, along with phase-number uncertainty relation, gives the uncertainty in phase as $\Delta\phi = 1/2\langle n \rangle^{1/2}$. This last expression is the starting point adopted by Dorschner et al. Each mode of the cavity is in a coherent state with cavity lifetime τ. Let the initial and final phases of a state be ϕ_i and ϕ_f respectively so that the frequency measurement of the mode in a single lifetime gives $\omega = (\phi_f - \phi_i)/\tau$. The uncertainty in the measurement of ω is expressed as $\Delta\omega = (\Delta(\phi_f - \phi_i))/\tau$. Any real measurement of frequency will be performed over a measurement time T which is much greater than the coherence time τ; this is written $T = N\tau$ so that

$$\begin{aligned} \Delta\omega &= \frac{1}{T}\Delta(\phi(T) - \phi(0)) \\ \Delta\omega &= \frac{1}{T}\Delta([\phi(T) - \phi(T-\tau)] + [\phi(T-\tau) - \phi(T-2\tau)] + \cdots \\ &\quad \cdots + [\phi(\tau) - \phi(0)]) . \end{aligned} \tag{4.7}$$

Each of the pairs in square brackets are independent and identically distributed so that the total uncertainty is equal to the uncertainty of a single pair multiplied by $\sqrt{N}$. Further, the uncertainty in the remaining pair is $\sqrt{2}\Delta\phi$ so that

$$\Delta\omega = \frac{1}{\sqrt{2T\tau\langle n \rangle}} \tag{4.8}$$

and as might be expected a longer measurement time gives a more accurate value for the frequency. The linewidths and fluctuations in output frequency are often expressed in terms of the power injected per mode, P, and cavity quality factor, Q. Following Dorschner et al. τ and $\langle n \rangle$ may be eliminated using the relations $\tau = Q/\omega_0$ and $\langle n \rangle = \tau P/\hbar\omega_0$. The frequency fluctuations are then given by

$$\Delta\omega_{\text{rms}} = \sqrt{\frac{\hbar\omega_0^3}{2TPQ^2}} . \tag{4.9}$$

The relationship between these rms frequency fluctuations and the linewidth of the output beat frequency is quoted as $\Gamma_{\text{rad}} = T\Delta\omega_{\text{rms}}^2$ (see for example Statz et al. [107, pp.311]). This gives a linewidth which is independent of measurement time

$$\Gamma_{\text{rad}} = \frac{\hbar\omega_0^3}{2PQ^2} \quad \text{or} \quad \Gamma_{\text{Hz}} = \frac{1}{2\pi}\frac{\hbar\omega_0^3}{2PQ^2} . \tag{4.10}$$

It is this expression of linewidth that will be compared with other derivations in §4.2.2. The independence of this linewidth from the measurement time is expected as it is a fundamental property of the mode. However, the accuracy with which the line centre may be located improves as $T^{-1/2}$.

4.2.2 Alternative Derivations of Quantum Noise Induced Linewidth

Two further derivations of the quantum noise induced linewidth are examined here. These derivations give results which appear to differ from that calculated above. Each derivation includes a term resulting from the specifics of the laser operation and which, when properly suppressed, reduces to the same limit as above.

The first derivation considered is that due to Sargent, Scully and Lamb [101] which is also adopted by Goldberg et al. [53] and Wilkinson [126]. These authors consider a fully quantised system, that is, both the optical field and the atoms in the gain medium are treated quantum mechanically. In this way extra features of the atoms that provide gain are included, in particular, a population inversion term is added to the final linewidth. In this analysis a differential equation [101, (p.334 Eq(48))] is found that governs the behaviour of the field amplitude $A(t)$. This amplitude is approximated by the steady state solution $A(t) = \bar{n}_{ss}^{1/2} e^{-i\phi(t)}$ where the optical frequency time dependence has been suppressed. The spectrum is expressed as

$$I(\omega) = \int_{-\infty}^{\infty} \exp(i\omega t) \langle \boldsymbol{A}^{\dagger}(t) \boldsymbol{A}(0) \rangle dt \tag{4.11}$$

and a phasor analysis provides $\langle \boldsymbol{A}^{\dagger}(t) \boldsymbol{A}(0) \rangle = \bar{n}_{ss} \exp(-\langle D(\phi) \rangle t)$. From consideration of the spectrum the linewidth (FWHM), $2\langle D(\phi) \rangle$, is found to be

$$\Gamma_{\text{rad}} = 2\langle D(\phi) \rangle = \frac{1}{2} \frac{\nu}{Q \bar{n}_{ss}} (\bar{n} + |\bar{n}_m|) \tag{4.12}$$

where $D(\phi)$ is the diffusion constant of the phasor $e^{-i\phi(t)}$ and $\bar{n}$ is the thermal contribution which is shown by the authors to be negligible. In order to compare this result with other linewidth calculations $\bar{n}_{ss}$ needs to be eliminated. This is achieved by considering the definition of cavity Q which is

$$Q = 2\pi \frac{\text{Energy stored in cavity}}{\text{Energy loss per cycle}} = \frac{2\pi \bar{n}_{ss} \hbar \omega}{P/\omega} = \frac{2\pi \bar{n}_{ss} \hbar \omega^2}{P} \,. \tag{4.13}$$

Some conventions do not include the factor of 2π and hence care has to be taken when comparing the results of different authors. Eliminating $\bar{n}_{ss}$ and noting that $|\bar{n}_m| = N_a/(N_a - N_b)$, in their notation, this gives

$$\Gamma_{\text{rad}} = \frac{\hbar \omega^3}{2PQ^2} \left(\frac{1}{1 - g_b n_b / g_a n_a} \right) \tag{4.14}$$

which is to be compared with Eq. (4.10). In this expression n_a, g_a and n_b, g_b denote to the populations and degeneracies of the upper and lower lasing levels respectively. Dorschner et al. [41] suggest that this factor is non-negligible and may give a correcting factor as large as 0.4-0.6.

A further method for considering phase fluctuations is described by Cresser and coworkers in a series of papers [33–36] and considers the time evolution of the phase difference between the two counter-propagating modes in terms of stochastic differential equations. A review by Chow et al. [28, §V] follows Cresser et al. and concisely summarises the main results. The techniques used to solve the stochastic differential equations are highly sophisticated; only an outline of the process is provided here.

The stochastic differential equation that governs the phase evolution of ϕ, the phase difference between the two modes, in the absence of any mode coupling is written

$$\dot{\phi} = a + F(t) \; . \tag{4.15}$$

Such stochastic equations are called Langevin equations. The uncoupled output frequency is a and the term $F(t)$ represents a stationary fluctuating contribution, $\langle F(t) \rangle = 0$. The solution to this differential equation is expressed as a power spectrum

$$\alpha(\omega) = \frac{1}{2}\left(\frac{D}{(\omega - a)^2 + D^2} + \frac{D}{(\omega + a)^2 + D^2} \right) \; . \tag{4.16}$$

The diffusion constant, D, is quoted by Chow et al. as $\omega/2Q\langle n \rangle$ which leads to a linewidth twice that found earlier in Eq. (4.10). The extra factor of two is traced back to the quoted value for D. Further information is provided here in that the power spectrum has a Lorentzian profile. These stochastic calculations provide somewhat more interesting information than simply to confirm the results found from other methods. In particular, the power spectrum in the presence of Adler-like coupling between the modes is informative. It is known that this kind of coupling gives a lock-in band around the zero frequency separation region; the precise behaviour, in the presence of a white frequency noise term, is found. Cresser et al. find exact solutions, using infinite continued fractions [98], to the stochastic differential equation

$$\dot{\phi} = a + b \sin \phi + F(t) \tag{4.17}$$

making use of the Fokker-Planck method. This coupling between the two modes is expressed in the same manner as the Adler equation for two coupled oscillators and has coupling coefficient b. An approximate analytic result is found which is valid in the case of weak noise, $b \gg D$,

$$\begin{aligned} \alpha(\omega) \;=\; & \frac{\pi D^2}{2}\left(\frac{|a| - \Gamma}{b} \right)^2 \left(\frac{|a| - \Gamma}{|\lambda_2|^2} - \frac{2|a|}{|\lambda_1|^2} \right)^2 \delta(\omega) \\ & + \left(\frac{\Gamma}{b} \right) \sum_{n=-\infty, \neq 0}^{\infty} \left(\frac{|a| - \Gamma}{b} \right)^{2|n|} \frac{2n^2 d}{(\omega - n\Gamma)^2 + n^4 d^2} \end{aligned}$$

where

$$\begin{aligned} \Gamma &= (a^2 - b^2)^{1/2} \\ d &= \frac{D(a^2 + \frac{1}{2}b^2)}{\Gamma^2} \\ \lambda_n &= -in\Gamma + n^2 d \, . \end{aligned}$$

It can be seen that the fundamental and its harmonics are Lorentzian. It is also found [35] that the $\delta(\omega)$ part of the spectrum (the first term) falls quickly to zero as the modes unlock. The width of these spectral lines are dependent on the strength of the noise, D, the order of the harmonic, n, and the coupling between the modes, b:

$$\begin{aligned} \Gamma_{\text{rad}} &= \frac{2n^2 D(a^2 + \frac{1}{2}b^2)}{a^2 - b^2} \\ &\simeq 2n^2 D\left(1 + \frac{3b^2}{2a^2}\right) \quad \text{when} \quad b \ll a \, . \end{aligned} \tag{4.18}$$

The linewidth induced by quantum noise remains equal to $2D$ but the overall linewidth is increased from this limit due to the mode coupling. Also the linewidth increases quadratically with the order of the harmonic so that minimum linewidth is only found for the fundamental.

The solutions to the Adler equation, $\dot{\phi} = a + b\sin\phi$, are quoted inside and outside the lock-in region separately. Inside the lock-in region, $|a| \leq b$, and the mean output frequency (or time averaged output frequency) is zero, $\langle\dot{\phi}\rangle = 0$. When unlocked, $|a| > b$, the mean output frequency is $\langle\dot{\phi}\rangle = (a^2 - b^2)^{1/2}$. For a given coupling constant, b, the lock-in band is located between $\pm b$ with an abrupt change (discontinuous slope) in the dependence of the mean output frequency on the beat frequency, a, at these points. Abrupt changes of this kind are not usually found in classical systems. It is not surprising, then, that once a noise source is added to the Langevin equation this discontinuous behaviour vanishes. Cresser et al. [35] have evaluated the time averaged beat frequency as a function of a for a given coupling constant; their numerical results are shown in Fig. 4.2. The comparison between the noise-free case and two different noise levels is shown and it is found that the stronger the noise the further the predicted curve departs from the noise-free curve. An intuitive feel for these results may be obtained by considering two noise-free oscillators which are locked together but are near the unlocked region. In this case only a small additional separation between the modes will unlock them. This additional separation is provided by the noise term which, although a centre zero term, may fluctuate on the short time scale to unlock the modes momentarily. As the noise term is increased in amplitude the oscillators overcome their tendency to lock more easily and oscillate at a higher frequency.

Presented here and in §4.2.1 are three different methods of acquiring a fundamental limit for the resolution attainable by a ring laser gyro. The key point to be made is that all approaches give the same quantum limit once the effects

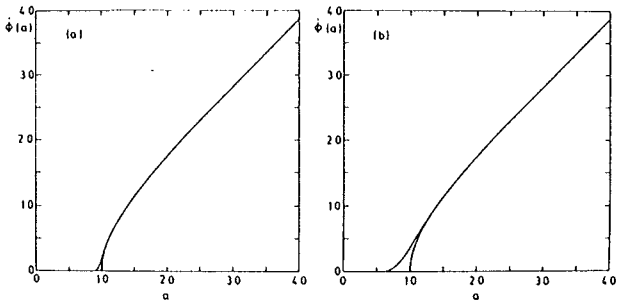


FIG. 5. Mean beat frequency $\langle\dot{\phi}\rangle$ as a function of the rotation rate a, compared to the noiseless response curve, $b=1$. (a) $D=10^{-2}$ and (b) $D=10^{-1}$.

Fig. 4.2: The numerical results of Cresser et al. [35]. Note that with a greater noise contribution the oscillators unlock more easily. For oscillators that are significantly far from the lock-in region the behaviour tends towards the Adler solution and $\langle\dot{\phi}\rangle = (a^2 - b^2)^{1/2}$.

of population inversion and mode coupling have been suppressed by appropriate design. All the Canterbury ring lasers use super-mirrors and hence are high Q cavities. In the first case this means that in single mode operation the gain per pass is extremely small and the population inversion correction term is very near to unity. In the second case our lock-in thresholds are typically much smaller that the Sagnac frequency and $b \ll a$ is a valid approximation. This implies that the correction term in brackets in Eq. (4.18) is also very near unity. Therefore under the operating conditions found at Canterbury all analyses converge on the same expression for linewidth, namely that given in Eq. (4.10).

4.2.3 Surpassing the Quantum Limit

The name of this section is possibly misleading in that the quantum limit that is being surpassed is only the limit that has been quoted as standard in much of the gyro literature. Recent developments in the description of quantum phase are discussed in §4.3 and provide the potential for the above quantum limits to be bettered in real systems. The attainment of these new quantum limits is not part of the local project, however, a brief discussion of what current theory now regards as the quantum limit is included for completeness and interest.

Minimum uncertainty states, such as coherent states, are those that satisfy the relation $\mathrm{var}(p)\mathrm{var}(x) = 1/4$ where $\hat{x} = (\hat{a} + \hat{a}^\dagger)/\sqrt{2}$ and $\hat{p} = (\hat{a} - \hat{a}^\dagger)/i\sqrt{2}$. The limit placed on the accuracy with which phase may be determined is given by the number-phase uncertainty relations which are often expressed as a power law of the form $\Delta\phi \sim \langle n\rangle^k$ where the standard quantum limit is given when $k = -1/2$. The key to reducing this limit lies in the use of squeezed states; a digestible primer to this topic is given by Henry and Glotzer [58]. Squeezed states have the property that their uncertainty product oscillates and that they

satisfy the minimum uncertainty relation 4 times each cycle. Fig. 4.3 illustrates the way in which Wigner contours are represented in position-momentum space and in number-phase space; these ideas are discussed in §4.3. A Wigner contour subtends an angle, $\Delta\phi$, at the origin and this angle is associated with the phase uncertainty of the mode. For a coherent state the Wigner function has circular contours and hence subtends the same angle wherever it is located on the circular dynamical path. The Wigner contour for squeezed states is not circular and under the proper circumstances may subtend a smaller angle at the origin than a coherent state. The uncertainty $\Delta\phi$ is now dependent on the orientation of the contour relative to the vector from the origin. At different parts of the cycle the Wigner function appears narrower than at others. If the measurement is taken at the proper part of the cycle then the observed uncertainty in the phase is below the quantum limit quoted for a coherent state. This is a very simple picture but the essentials depicted here remain valid. The details of phase space calculations are not investigated further.

The power law that relates phase uncertainty to expected photon number for squeezed states has been shown by D'ariano et al. [38, 39] to be $\Delta\phi \sim \langle n\rangle^{-1}$. These authors propose a feasible means by which this limit may be realized and that the precise relationship is $\Delta\phi \simeq 1.36/\bar{n}$. Further, it is shown that the exponent in this power law is only weakly dependent on the non-unit quantum efficiency of the photo-detectors. Such power laws are only valid for sufficiently large $\langle n\rangle$ because $\Delta\phi$ is bounded by $\pm\pi$ (see Opatrný [84]). Although the idea of a ring laser that operates with squeezed states is not under consideration it remains interesting to see what the quantum limits of such a device would be. Repeating the calculation that leads up to Eq. (4.10) the FWHM linewidth is found to be $\hbar^2\omega_0^5/2Q^3P^2$ representing a reduction of $\hbar\omega_0/QP$. This brings the quantum limit for one of the Canterbury Ring Laser group's ring lasers, say C-I, down by a factor of $\sim 10^7$. Improvements of fundamental limits by this sort of factor explain the current interest in squeezed states of light.

4.3 Quantum Phase

The problem behind the debate regarding quantum phase can be easily stated; what quantum measurable corresponds to the classically well understood concept of the phase of an electromagnetic field? This debate has been unresolved since it started at the birth of quantum mechanics. The debate is relevant to the ring laser project as the low light levels that are associated with larger ring lasers may ultimately require a more sophisticated description of phase. For the ring lasers constructed to date it is shown here that the semi-classical description of phase (summarised by the relation $\Delta\phi\Delta n \geq 1/2$) is sufficient to describe the system. However, as ring lasers become larger and the free spectral range becomes smaller the method of starving all but a single mode of the necessary gain required to lase will ultimately lead to light levels so low that the semi-classical description of phase fails. This light level is found to be of the order of a few pW which is

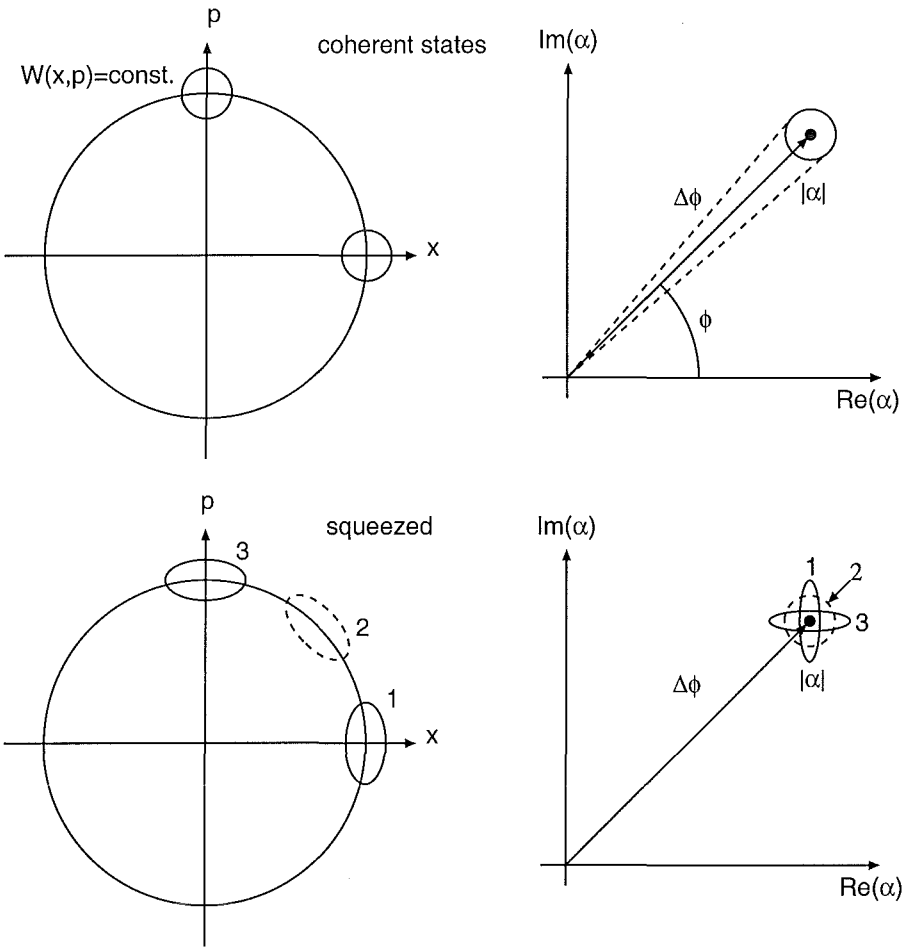


Fig. 4.3: An illustration of the Wigner contours in phase space for coherent and squeezed states. The figure is a composite from similar figures appearing in Henry et al. [58] and Lynch [74]. The upper two plots show the appearance of a coherent state in position-momentum space and α-space. Similar plots are shown beneath for a squeezed state.

not far below the output power of C-II. An introduction to this topic is provided in an excellent and up-to-date (as of mid 1995) review by Lynch [74]. A more recent tutorial review is provided by Pegg and Barnett [89] which provides a roughly chronological list of references regarding quantum phase. This list is quite comprehensive and includes over 450 references. The background of the topic and review of the literature is provided here.

4.3.1 Introduction to Quantum Phase

The classical concept of phase has historically been the starting point of quantum phase descriptions. This approach starts by considering a cubic cavity with edge L. The discussion is restricted to a single excited mode of the cavity with a

particular polarisation

$$E(\boldsymbol{r},t) = \left(\frac{\hbar\omega}{2\epsilon_0 L^3}\right)^{\frac{1}{2}} [ae^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} + a^* e^{-i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}]$$

where a is the complex field amplitude and is equated with $Re^{i\phi}$, comprising a real amplitude R and a phase ϕ. Thus

$$E(\boldsymbol{r},t) = E_0 \cos(\boldsymbol{k}\cdot\boldsymbol{r} - \omega t + \phi) \ ,$$

where $E_0 = (2\hbar\omega/\epsilon_0 L^3)^{1/2} R$. This is the manner in which the classical field is decomposed into an amplitude and a phase. There are several different approaches that may be followed in order to arrive at a viable theory of how this classical amplitude and phase picture may carry over to the quantum domain. Here I will illustrate the first and most natural approach which was formulated by Dirac [40]. The canonical quantisation procedure requires that the classical field be written in terms of generalised position and momentum coordinates, q and p, and these variables are related to the quantum mechanical operators $\hat{q}$ and $\hat{p}$. These conjugate operators are defined in terms of the creation and annihilation operators according to

$$\hat{q} = \left(\frac{\hbar}{2\omega}\right)^{\frac{1}{2}} (\hat{a} + \hat{a}^\dagger) \quad \text{and} \quad \hat{p} = -i\left(\frac{\hbar\omega}{2}\right)^{\frac{1}{2}} (\hat{a} - \hat{a}^\dagger) \ .$$

The electric field operator may then be written in a form analogous to the classical case given above

$$\hat{\boldsymbol{E}}(\boldsymbol{r},t) = \left(\frac{\hbar\omega}{2\epsilon_0 L^3}\right)^{\frac{1}{2}} [\hat{a}e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)} + \hat{a}^\dagger e^{-i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}] \ .$$

In quantising phase Dirac defined the two Hermitian operators $\hat{R}$ and $\hat{\phi}$ according to the relations

$$\hat{a} = e^{i\hat{\phi}}\hat{R} \quad \text{and} \quad \hat{a}^\dagger = \hat{R}e^{-i\hat{\phi}} \ .$$

The form of $\hat{R}$ can be found considering $\hat{a}^\dagger\hat{a} = \hat{n} = \hat{R}^2$ giving $\hat{R} = \hat{n}^{1/2}$ and hence the phase operator is found $e^{i\hat{\phi}} = \hat{a}\hat{n}^{-1/2}$. From these relations it is possible to use the identity $[\hat{a}, \hat{a}^\dagger] = 1$ and the fact that $\hat{n}^{-1/2}$ and $\hat{n}$ commute to derive the result $[e^{i\hat{\phi}}, \hat{n}] = e^{i\hat{\phi}}$. Expanding the exponential on both sides of this equation and equating terms of lowest order the commutation relation between the number operator and the phase operator is found $[\hat{n}, \hat{\phi}] = i$. This commutation relation leads to the uncertainty relation $\Delta n \Delta\phi \geq 1/2$ which is used later in this chapter (see §4.2.1). Unfortunately this rather natural and logical line of reasoning is incorrect. This is most clearly seen by considering the matrix elements of the commutator in the number state basis

$$\begin{aligned} \langle n|\hat{n}\hat{\phi} - \hat{\phi}\hat{n}|m\rangle &= i\langle n|m\rangle \\ (n-m)\langle n|\hat{\phi}|m\rangle &= i\delta_{n,m} \ . \end{aligned}$$

When $n = m$ a contradiction is found and hence this method of handling quantum phase fails. This sequence of steps is included to illustrate that the problem is real. The straight-forward approach fails to give a self consistent theory of how phase is to be described at the quantum level[1].

4.3.2 Alternative Theories for Quantum Phase

The problem outlined above can be corrected by extending the number states from ∞ to $-\infty$ (see Barnett and Pegg [14]) as a formal construct. However, the formalism is beset by more fundamental problems. A series of operator theories have been proposed each of which deals with the problems left by the last. Most notable among these theories, in chronological order, are those due to Susskind and Glogower [113], Garrison and Wong [48], Paul [87] and Lévy-Leblond [71]. This last paper proposes among other things the radical idea of abandoning Hermiticity as a requirement of all real operators which is a massive departure from convention. This amounts to a challenge to the axioms of quantum mechanics and highlights the importance of acquiring a proper description of phase.

A separate line of enquiry is to consider a system defined by a finite number-state space. The form of the phase operator in such a space is quite clearly defined and also has a discrete spectrum. The dimension of the system is ultimately taken to infinity in order to model a continuous phase operator. The theory of Pegg and Barnett [88] is widely employed. These authors carry the finite dimensional nature through their calculations only taking the limit to infinity when all the physical quantities have been found. Popov and Yarunin [93] also follow this line of reasoning.

If a system is described in terms of a coherent state $|\alpha\rangle$ then a phase space description of the mode may be used with the real and imaginary parts of α locating a point in this two-dimensional space that represents the mode. The coherent state is defined as the particular weighted sum over number states given in Eq.(4.2.1) where α is any complex number. A standard result says that $\langle\alpha|\hat{n}|\alpha\rangle = |\alpha|^2$ which suggests that $|\alpha|$ may be associated with an amplitude term suggesting the separation of amplitude and phase terms according to $\alpha = |\alpha|\exp(i\phi)$. In this way a phase space description gives an alternative, non-operator-based approach to quantum phase. In order to gain a qualitative idea about how this phase space description operates a probability distribution $W(\alpha, \alpha^*)$ is defined. Often this is the Wigner distribution which is used in quantum mechanics to define a probability distribution for position and momentum $W(q, p)$. The picture out-

[1] All the theories mentioned here, in the article by Lynch [74] and the thesis by Tsui [116] are consistent at the macroscopic level. For a summary of how different theories match up when predicting fluctuations see Tsui [116, Fig. 4.2] and Lynch [74, Figs. 12.2, 13.3 and 13.4] or Figs. 4.4, 4.6 and 4.7 in this chapter. It is worth noting here that all the theories converge to the same predictions of phase uncertainty for measurements that involve a dozen or more photons and by the time the photon count is up to 30 or more the system behaves according to the number-phase uncertainty relation.

lined here uses the idea of probability contours defined by $W(\alpha, \alpha^*) = \text{constant}$ (see Fig. 4.3). The uncertainty in the phase is defined as the arc subtended by the appropriate contour. Clearly this is an oversimplified description but is, in essence, valid.

The method proceeds by considering a phase-eigenstate $|\phi\rangle$ and forming the probability density $P(\phi)$ according to $|\langle\phi|\psi\rangle|^2$ where $|\psi\rangle$ is the state that describes the system under investigation. This probability density may be found via a variety of different methods. One method due to Schleich et al. [102] evaluates the overlap area of $|\phi\rangle$ with $|\psi\rangle$ where the phase eigenstate is taken to be a wedge originating from the origin; this is a valid approximation for large $|\alpha|$. Further phase space descriptions are offered by Vogel and Schleich [121] and by Freyberger and Schleich [46]. A detailed discussion of these methods is not pursued here as it is not within the bounds of the thesis topic. However a comparison with the experimental evidence is relevant to the experiment proposed in §4.3.5 and is provided later in Fig. 4.7.

4.3.3 Quantum Phase Experiments

As can be seen from the above comments a proliferation of theories governing quantum phase has arisen in the last few decades. However, until recently, experimental feedback has not been able to differentiate between these theories. The main experimental data for much of this period has been that due to Gerhardt et al. [49], the details of which are not elaborated here. This data is illustrated in Fig. 4.4 along with the predictions of several of the more highly regarded quantum phase theories. Clearly the experimental data is limited. There are only four data points and the uncertainties are sufficiently large that none of the theories shown can be ruled out. Note that at large n all the theories converge to the theory which is derived from the number-phase uncertainty relation $\Delta\phi\Delta n \geq 1/2$, where the equality holds for a coherent state such as that considered here. It follows that $(\Delta\phi)^2(\Delta n)^2 = 1/4$ so that $(\Delta\phi)^2 = 1/2n$ where the extra factor of 2 arises from the Gerhardt et al. experiment which considers the difference in phase between two beams.

A revolution in quantum phase theory has occurred recently with the experimental results of Noh, Fougères and Mandel [81–83] (sometimes written NFM in the literature). These authors have performed an excellent experiment, a schematic for which is provided in Fig. 4.5. The principle used is that of optical homodyne detection which is similar to heterodyne methods except that the two beams are of the same frequency. This is achieved by deriving both beams from the same highly stabilised laser output. The arrangement used by NFM (Fig. 4.5 (b)) is somewhat more complex than the usual optical homodyne system illustrated in Fig. 4.5 (a). This complexity is required to give simultaneous measurements of the sine and cosine of the phase difference between the two fields.

The theory provided by NFM is their own and is defined operationally, that is,

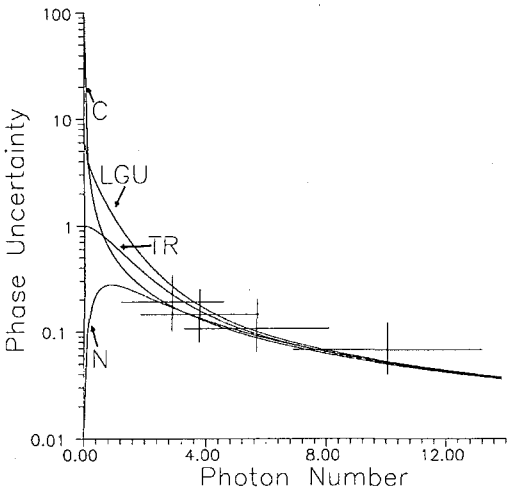


Fig. 4.4: This figure illustrates the predictions of various authors of the phase uncertainty $(\Delta\phi)^2$ against photon number n. The figure is taken from Tsui and Reid [115] and the data points are from Gerhardt et al. [49]. The curve labelled TR is from the theory of Tsui and Reid. The standard analysis of phase, ostensibly Dirac's, is labelled C. The predictions of Nieto [80] are labelled by N and follows the Susskind and Glogower [113] formalism. The last curve is from Lynch [73] and Gerry and Urbanski [50], labelled LGU, and follows the Pegg and Barnett [88] approach.

specifically in terms of the experimental arrangement used. The authors define operators which are related to the cosine and sine of the classical phase difference between the input fields. The operators so defined are particular to this experiment and do not provide a global description of phase. The classical representation of the fields is quantised and yields the expressions

$$\begin{aligned} \hat{n}_4 - \hat{n}_3 &= \tfrac{1}{2}[(\hat{a}_1^\dagger - i\hat{a}_{10}^\dagger)(\hat{a}_2 - i\hat{a}_{20}) + \text{H.c.}] \equiv \hat{C} \\ \hat{n}_6 - \hat{n}_5 &= \tfrac{1}{2}[(-i\hat{a}_1^\dagger + \hat{a}_{10}^\dagger)(\hat{a}_2 + i\hat{a}_{20}) + \text{H.c.}] \equiv \hat{S} \end{aligned}$$

where the subscripts 10 and 20 refer to the vacuum fields (see Fig. 4.5), the subscripts 1 and 2 represent the input fields and the subscripts 3 to 6 represent the fields at the detectors. It is the association of the $\hat{C}$ and $\hat{S}$ operators with the cosine and sine of the phase difference between the input fields that gives meaning to quantum phase in this experiment. These operators have the property of commuting with one another which is desirable as the measurements are made simultaneously. The operators require a rescaling in order to be compatible with the experimental output. This is achieved according to the relations

$$\begin{aligned} \hat{C}_T &= \tfrac{1}{2}[(\hat{a}_1^\dagger - i\hat{a}_{10}^\dagger)(\hat{a}_2 - i\hat{a}_{20}) + \text{H.c.}](\hat{C}^2 + \hat{S}^2)^{-1/2} \\ \hat{S}_T &= \tfrac{1}{2}[(-i\hat{a}_1^\dagger + \hat{a}_{10}^\dagger)(\hat{a}_2 + i\hat{a}_{20}) + \text{H.c.}](\hat{C}^2 + \hat{S}^2)^{-1/2} \, . \end{aligned}$$

These final operators also commute with each other and also with the total photon number $\hat{n}_1 + \hat{n}_2 + \hat{n}_{10} + \hat{n}_{20}$. This theoretical description is compared with the following experimental interpretation. NFM define a *measured* sine, S_M, and cosine, C_M, so that if each sample of the system delivers the four integers

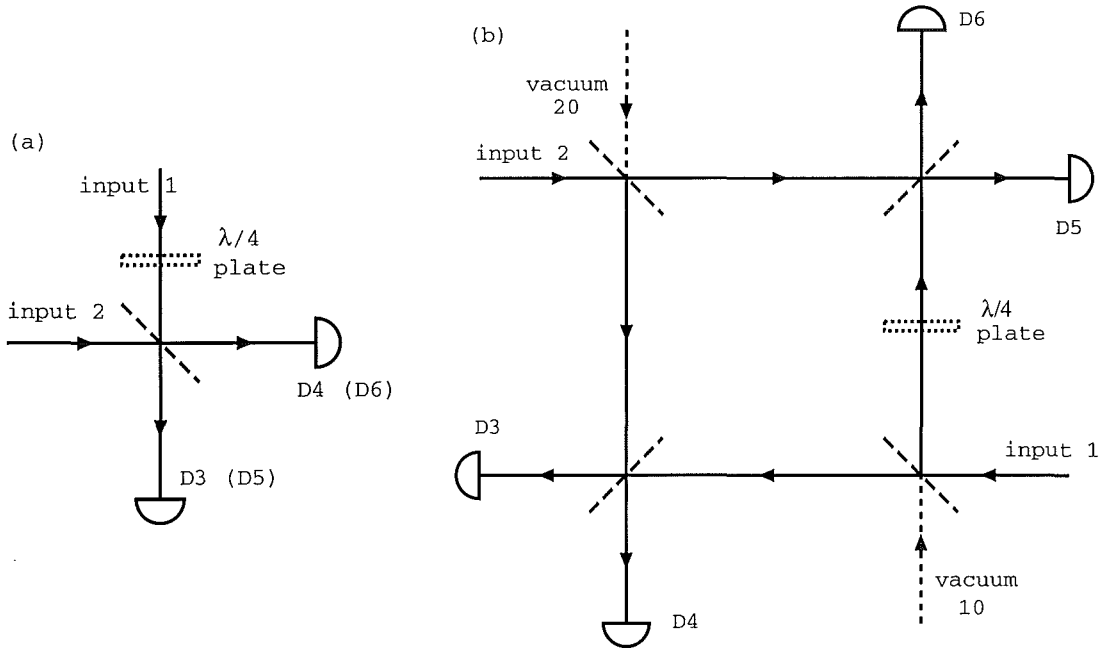


Fig. 4.5: These schemes are taken from Noh et al. [81]. Scheme (a) on the left shows the basic homodyne detection arrangement. When the quarter-wave plate is absent the measurements taken by detectors D3 and D4 are used to evaluate the cosine of the phase difference between the two input fields. Inserting the plate allows a measurement of the sine of this angle using the counts from D5 and D6. Scheme (a) has only limited application and all the results that are illustrated later are from Scheme (b) on the right. This scheme allows the simultaneous measurement of the sine and cosine of the phase angle between the two input fields.

(m_3, m_4, m_5, m_6), corresponding to photon detections, then

$$C_M = \frac{m_4 - m_3}{((m_4 - m_3)^2 + (m_6 - m_5)^2)^{1/2}}$$
$$S_M = \frac{m_6 - m_5}{((m_4 - m_3)^2 + (m_6 - m_5)^2)^{1/2}} .$$

The samples where $m_3 = m_4$ and those where $m_5 = m_6$ are discarded. In practice the joint probability distribution $P(m_3, m_4, m_5, m_6)$ is found by taking a great many samples. The instances where $m_3 = m_4$ or $m_5 = m_6$ still provide no information about phase and are removed from the sample space. The removal of these samples requires a renormalisation of the joint probability distribution. Further, the effects of background counts have to be corrected. These fine points are covered in great detail in the papers by NFM and are not repeated here. The main result is that once a corrected probability distribution, P_c, is found the moments of the measured sine and cosine of the phase difference may be found

according to

$$\langle \cos_M^r(\phi_2 - \phi_1)\rangle = \sum_{\{m\}} \frac{(m_4 - m_3)^r}{((m_4 - m_3)^2 + (m_6 - m_5)^2)^{r/2}} P_c(m_3, m_4, m_5, m_6)$$

$$\langle \sin_M^r(\phi_2 - \phi_1)\rangle = \sum_{\{m\}} \frac{(m_6 - m_5)^r}{((m_4 - m_3)^2 + (m_6 - m_5)^2)^{r/2}} P_c(m_3, m_4, m_5, m_6)$$

The extraordinary results that have been found by this experiment are reproduced in Fig. 4.6. Implicit in this figure is the assumption that $\langle(\Delta\hat{C}_M)^2\rangle + \langle(\Delta\hat{S}_M)^2\rangle \approx \langle(\Delta(\phi_2 - \phi_1))^2\rangle$. The theory of Susskind and Glogower [113] departs markedly from the experimental data but only when the average number of photons is small. However, the selection of the theory presented by NFM over that presented by

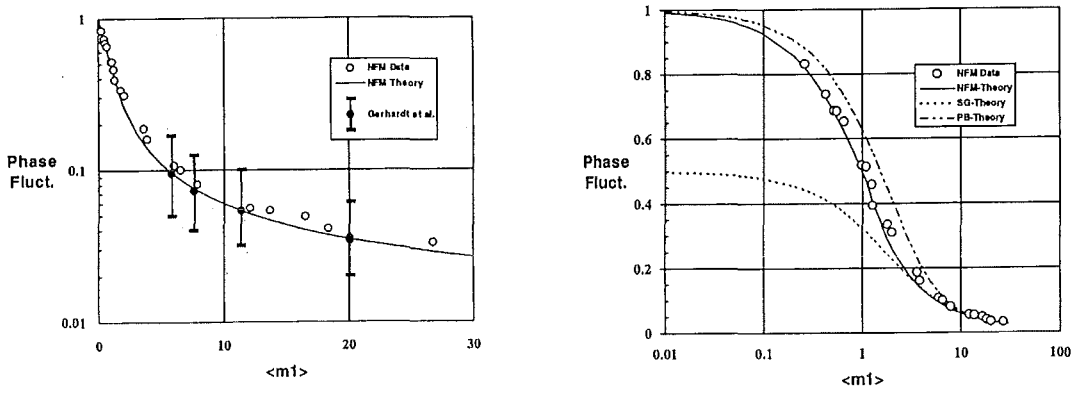


Fig. 4.6: Both of these figures are taken from Lynch [74]. The figure shows how the NFM data [81] compares with that taken by Gerhardt et al. [49] and suggests that these latter authors have overestimated there uncertainties. On the right hand side the experimental data and theoretical predictions of NFM are presented alongside the predictions of Pegg and Barnett [88] and Susskind and Glogower [113]. The winning theory is clear.

Pegg and Barnett [88] is only possible due to the very low uncertainties afforded by this experiment which are smaller than the open circles on the plot.

There remain reservations about the interpretation made by NFM regarding their experiment. The idea that a phase description is dependent on the particular arrangement under consideration has had some authors looking for alternative explanations of the NFM data. The main results of these efforts are shown on Fig. 4.7 where it can be seen that the Vogel-Schleich [121] theory fails for small photon numbers but the Freyberger-Schleich [46] theory agrees nearly perfectly with the NFM data.

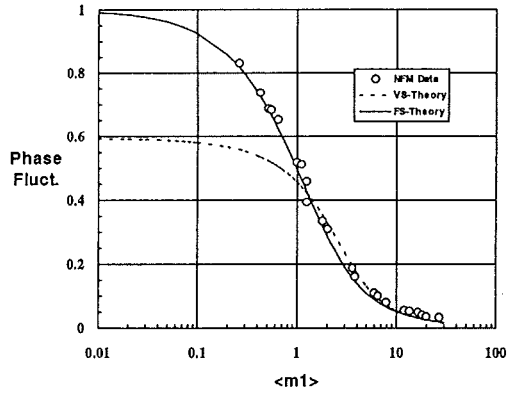


Fig. 4.7: This figure is taken from Lynch [74]. The theoretical curve that does not fit the data for low photon numbers is due to Vogel and Schleich [121] whereas the other curve, which fits the data almost perfectly, is due to Freyberger and Schleich [46].

4.3.4 Relevance of Quantum Phase to the Local Project

In the previous section it has been shown that all the available theories of quantum phase converge for $n \gtrsim 12$. When a particular measurement can be traced to a number of photons greater than 12 the use of the number-phase uncertainty relation is valid to good approximation. To get an idea whether a sophisticated quantum phase description is required in any of the ring laser experiments performed until now a few order-of-magnitude calculations are performed. C-I has been operated at 500 nW which corresponds to a photon flux of $\sim 1.7 \times 10^9\,\mathrm{s}^{-1}$ (where $\hbar\omega \approx 3 \times 10^{-19}\,\mathrm{J}$). Although the photomultiplier tubes used as detectors have bandwidths of about 700 MHz the overall bandwidth of the detection system is governed by the subsequent electronics especially the low-pass filters that are used to reduce the high frequency noise. The particular Hamamatsu PMTs used have been chosen for their good efficiency at 632.8 nm which is of the order of 10%. Typically the bandwidth of the detection system has been about 500 Hz and hence any measurement of the output from the PMT detectors is the result of $\sim 3.4 \times 10^5$ photons. Hence the observations made under these circumstances are properly described by the number-phase uncertainty relation. More recently the output power of C-II has been measured at 10 pW [108] which brings the total number of photons down to $\sim 7 \times 10^3$ assuming a similar detector bandwidth. Clearly this situation also falls within the scope of the uncertainty relation. In the next section it is found that under certain circumstances the number of photons involved in each measurement may fall below 12 in larger ring lasers. Under these conditions a more sophisticated phase description will be required to describe quantum noise.

4.3.5 An Experiment to Investigate Quantum Phase Using a Ring Laser

The previous section shows that the debate regarding the nature of phase at the quantum level does not appear to affect the experimental work undertaken so

far by the Canterbury ring laser group. However, it is not hard to attenuate the output beams and/or increase the bandwidth of the detection system to the point where the effective number of photons in each reading demands a more sophisticated description of phase. The attenuation of the beams to the appropriate level would allow the NFM experiment to be repeated by our group. To confirm this ground breaking work is worth while and, in addition, a further experiment suggests itself which could provide vital information about ring laser gyros.

The output beams of a ring laser gyro are usually interferometrically combined to give an oscillating output with frequency linearly proportional to the rotation rate. Initially one might think that the optical bandwidth of the cavity (10^3 to 10^5 Hz) suggests that the time dependent output, typically at audio frequencies, would be smeared out by the optical frequencies moving about under this bandwidth. This is not the case however as the *active* Q of the cavity is much larger than the passive Q and hence the active linewidth of the laser is vanishingly small. The lasing frequency is not stable however; even the best stabilised lasers still drift on the scale of kHz/s. How can independent oscillators, such as the counter-propagating modes of the ring laser, beat together to give such a stable audio tone? The only explanation is that the two optical frequencies drift stochastically but are locked together with a constant separation and, hence, are not truly independent. This is definitely not the case for two beams derived from separate lasers, which drift independently and make optical heterodyning a particularly difficult experiment to perform. This constant separation of frequencies, despite such large fluctuations in optical frequency, can only reasonably be expected to arise due to the two modes sharing the same cavity.

It is tempting to conclude that the fundamental limits for rotation detection by a ring laser gyro are contained in the number-phase uncertainty relation $\Delta n \Delta\phi \geq 1/2$ when the photon numbers are above a certain, very low, threshold. This conclusion, however, is made using the results of an experiment which considers the fluctuations in phase between two fields derived from the same laser source. The experiment proposed below uses the counter-propagating ring laser output beams as inputs to an arrangement which is essentially the same as that used by NFM. These beams differ from those used by NFM because they are not derived from the same source but are highly correlated due to their common geometry and sufficiently decoupled for the ring laser to unlock under the Earth's rotation. The results of such an experiment would be of great interest both to quantum phase theorists and to those interested in a specific test of the fundamental limits placed by quantum mechanics on the precision of ring laser gyros.

Such results would be of special interest to manufacturers of large, single mode gyros with extremely low cavity losses because such lasers necessarily have extremely low output powers. The increased perimeter reduces the free-spectral-range so that the gain needs to be reduced further to ensure single mode operation [99] and the small losses lower the gain required to ensure lasing. C-II has a typical output power of 10 pW and, as mentioned below, under the proper circumstances an output power of 3 pW may fall within the realm of a full quantum descrip-

tion of phase. This analysis may be particularly relevant to the proposed 16 m perimeter gyro 'G' being built (Jan 1998) for Bundesamt für Kartographie und Geodäsie, Technische Universität München and The University of Canterbury

A description of the quantum phase experiment is now provided with some detail regarding the required equipment and operating parameters. The overall

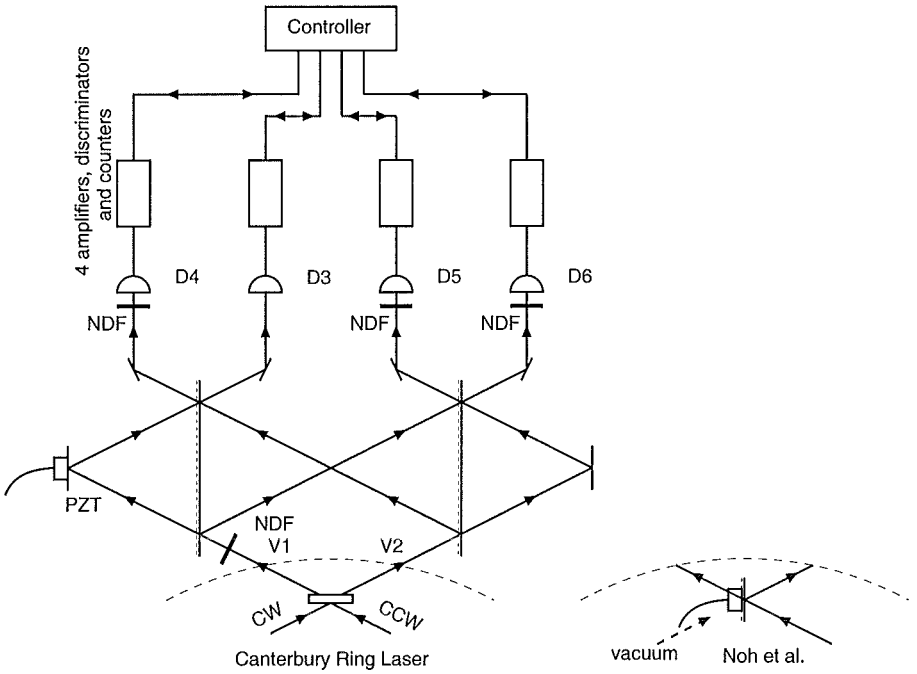


Fig. 4.8: This figure illustrates the proposed experiment. Above the dotted line the layout is the same as that of Noh et al. and below the line is the corner of one of our ring lasers; the corner mirror is shown but the mirror housing and body of the laser are not. The light source used by Noh et al. is included for completeness.

detection rate needs to be sufficiently low that the dead-time effects of whichever detector is used can be neglected. In contrast, this rate must also be sufficiently high that the dark current effects, usually quoted as a dark current induced count rate, can also be neglected. The typical pulse width is about 20 ns which places an upper limit on the count rate of about $10^7\,s^{-1}$. The dark current induced count rate for the appropriate photomultiplier at room temperature is of the order of $10^2\,s^{-1}$ which is required to be significantly smaller than the detection rate. This dark current rate may be reduced by a factor of about 3 by cooling if required but this is an involved procedure for a relatively small return. Hence a detection rate of about $10^4 - 10^5\,s^{-1}$ is required. With a typical quantum efficiency of $\sim 1\%$ (typical for PMTs with low dark current such as those used for photon counting) the incident power on the photomultiplier is required to be about 3 pW.

Fig. 4.8 illustrates the layout of the proposed experiment to be performed on one of our ring lasers. The original NFM experiment had a beam splitter located where the ring laser corner mirror is located on this diagram and is shown to the

right. The beam splitter had one input coupled to a stabilised He-Ne laser and the other coupled to the vacuum. The beam splitter was mounted on a piezoelectric transducer (PZT) to allow adjustment of the phase difference between the input fields. In the proposed optical heterodyne experiment this adjustment is not necessary when the ring laser is unlocked because the phase between the input fields, ϕ_0, is continuously evolving. Further, the relative phase between the input fields may be measured simultaneously at another port and from this ϕ_0 may be found. In this way the phase experiment may be performed with a variety of values of ϕ_0 simply by choosing when the measurement is performed. An alternative data collection method collects data continuously with a log of ϕ_0, deduced simultaneously from another exit port. This would require a different analysis than the optical homodyne case when ϕ_0 is constant. Data may be collected continuously, as suggested above, or synchronously so as to give constant ϕ_0. In either case the detection system is instructed to take samples by the controller; a computer fitted with the appropriate interface cards. The data is then relayed back to the computer and stored for later analysis.

The proposed experiment would extend the noise profile (discussed in §4.4) to short time scales. The uncertainty relation gives a -1/2 power law relationship between frequency fluctuations, Δf_{rms}, and measurement time, T, as seen on Allan deviation plots. This power law is expected to break down for small T reflecting the failure of the uncertainty relation to describe phase for low n. Agreement with the NFM experiments would require that for sufficiently small T the frequency fluctuations become independent of T. The benefit to the project comes in the form of greater knowledge of the frequency fluctuation spectrum.

4.3.6 Conclusion

The ongoing debate regarding quantum phase has a significant impact on the work undertaken by the Canterbury Ring Laser Group. The fundamental limit of the resolution of a gyro is defined by phase uncertainty which does not behave according to the number-phase uncertainty relationship when n is small. It has been shown (§4.3.4) that this does not affect the analyses that have been used as standard throughout the project but that future gyros may have sufficiently low light levels that demand a more sophisticated description of phase. To this end an experiment is proposed that uses an adaptation of the NFM experiment to investigate phase uncertainty when a ring laser's output is used as the light source. This experiment would operate in several different configurations. Initially a stabilised laser would be used to confirm the NFM results. Then the ring laser output would be substituted and the phase uncertainties explored when the laser is either locked or unlocked. This experiment differs from the NFM experiment because the two input beams are sourced from two non degenerate cavity modes. However, the modes share a common cavity and thus observation of interference between them is not difficult. This experimental work may provide a bridge between the foremost theory of quantum phase and state-of-the-art ring laser

gyros.

4.4 Noise Classification and Ring Laser Operation at the Quantum Limit

The stability of a ring laser is of utmost importance if it is to be a useful working device. In order to monitor the performance of a gyro a systematic method for displaying the stability as a function of integration time is required. Here the Allan variance is used for this role. Improvement in stability is logged by the lowering levels of frequency fluctuations as in Fig. 4.12. The different sources of noise give power laws which identify them clearly. The source of the noise is then more easily isolated and reduced. As one noise source is diminished another will become dominant and that in turn can be dealt with. The ultimate performance is limited by quantum noise at short time intervals and flicker noise (yet to be seen in any of the local ring lasers) at longer intervals. In this section the Allan variance method is documented especially with regard to ring laser gyro performance. It is shown that operation at short times is already at, or near, the quantum limit. At longer times environmental conditions prevail as the largest source of frequency fluctuations. These environmental variations can be handled in one of two ways. Either the design of the laser can be made insensitive to such variations (reflected in the progression from C-I to C-II) or the variations can be monitored and accounted for subsequently. This latter step is performed for data taken from C-II and a marked improvement in the stability is achieved.

4.4.1 Allan Variance

The characterisation of noise processes relevant to the ring laser is discussed. In particular noise types in the output beat frequency are identified. A discrete set of measurements of the output frequency, f_i $(i = 1, \ldots, N)$, is acquired each for a given measurement time τ_0 using the AR(2) method introduced in §3.6; zero dead time is assumed throughout. Conventional analysis of this data would involve finding the mean and standard deviation of the frequency estimates and quoting the uncertainty in the frequency being measured as one or two standard deviations either side of the mean. However, in the event that the frequency has $1/f$ noise (pink noise or flicker noise) the estimate for the standard deviation diverges with time. Another approach is needed for this kind of noise phenomenon. The presence of flicker noise may be identified in the Fourier Transform of the frequency measurements, however, investigations of the properties of stochastic processes by Barnes [13] and Allan [4][2] reveal a better method for classifying

[2] The introduction of this variance is due to Allan [4] who was interested in characterising the quality of time-frequency standards. When dealing with the best frequency standards the only method of investigating their performance is to beat two of them together. If the two clocks

noise types. This new method defines a new variance according to

$$\sigma_A^2 = \frac{1}{2(n-1)} \sum_{k=1}^{n-1} (f_{k+1} - f_k)^2 \tag{4.19}$$

for a set of n measurements and is known as the Allan variance. This definition is to be contrasted directly with the usual definition of variance

$$\sigma^2 = \frac{1}{n} \sum_{k=1}^{n} (f_k - \bar{f})^2 \tag{4.20}$$

where $\bar{f}$ is the mean of the set of measurements. The appearance of the factor of $1/2$ in the definition of the Allan variance may be seen to be consistent with the conventional definition of variance by considering the case where only white frequency noise is present. The σ_A^2 term is equal to the expectation value of $(f_{k+1}^2 - 2f_{k+1}f_k + f_k^2)/2$ which is simply $\langle f_k^2 \rangle$ if the measured frequencies are uncorrelated as in white frequency noise. In this case the Allan variance is equal to the standard variance.

The Allan variance is a function of the measurement time τ and is written $\sigma_A^2(\tau)$. The duration τ can be thought of as a characteristic time for a frequency measurement; frequency fluctuations that are long on the time scale of τ have little effect on the Allan variance estimate. The method for determining $\sigma_A^2(\tau)$ is to use the frequency measurements f_i, taken in time τ_0, and group n of them in clusters; the average of each cluster represents the frequency measurement taken in a duration $n\tau_0$. It is important that there is zero dead time between these measurements. The Allan variance (or Allan deviation) is then plotted on a log-log plot where the integer power laws become linear. There is clearly a link between the power spectrum, $S(f)$, of the frequency estimates and $\sigma_A(\tau)$. This link is given explicitly [79, Eq.(7)] by the expression

$$\sigma_A^2(\tau) = 4 \int_0^\infty S(f) \frac{\sin^4(\pi f \tau)}{(\pi f \tau)^2} \mathrm{d}f \tag{4.21}$$

and hence all the information contained in the power spectrum is also contained in the Allan variance.

There is, of course, uncertainty in the estimate of the Allan variance given by Eq. (4.19). The statistics of these estimates are discussed in a paper by Prestwich et al. [95] which derives the probability distribution for σ_A^2 for any number of samples. However for the purposes of classifying noise types in ring lasers a sufficient expression for the 1-σ uncertainty of the Allan variance is given by Ng and Pines [79] as

$$\% \text{ uncertainty} = \frac{100}{\sqrt{2(K-1)}} \tag{4.22}$$

are constructed identically and suitably decoupled then the statistics of the beat frequency are in principle defined by the statistics of only one clock.

where K is the number of clusters available for given τ. For a measurement set of a given length K becomes small for large τ and the uncertainty increases. It is common, therefore, to see error bars on plots that become large for long τ. These error bars are omitted from Fig. 4.12 in order to keep the plot uncluttered.

4.4.2 Noise Type Classification Using Allan Variance

The power of the Allan variance method lies in its ability to classify noise types. The paper by Ng and Pines [79] discusses the use of Allan variance to characterise ring laser gyro performance and is therefore of particular relevance. The types of noise that are considered here are white noise, $1/f$ flicker noise and random walk noise although other effects that have particular features under an Allan variance analysis are discussed. Any of the three above noise types may appear in either the frequency or the phase of the signal being measured. Thus a sinusoidal signal of the general form $I(t) \propto \sin(2\pi f(t) + \phi(t))$ is considered.

The method of classification is derived from the dependence of the Allan variance on the measurement time τ. All the noise types mentioned give rise to integer power law dependences of the Allan variance on τ. Different noise types give different power laws and thus can be distinguished from one another. When plotted on a log-log plot the power law dependences are linear and it is simple to distinguish the noise types; Table (4.1) tabulates their typical causes. Fig. 4.9 summarises the power law dependences for the various noise types. Some authors present results in terms of the Allan variance and others in terms of the Allan deviation; I have tended to discuss the normalised Allan deviation which scales $\sigma_A(\tau)$ down by the mean frequency and hence gives an absolute stability as a function of τ. In this way devices which have different operating frequency may be compared for absolute stability more easily.

The power law dependences may be verified using Eq. (4.21). For example, consider the two cases of white and flicker frequency noise; the Allan variance is evaluated to be:

$$\begin{aligned} \text{flicker frequency noise} \quad \sigma_A^2(\tau) &= 4\int_0^\infty \frac{1}{f}\frac{\sin^4(\pi f\tau)}{(\pi f\tau)^2}df = 4\ln(2)\tau^0 \\ \text{white frequency noise} \quad \sigma_A^2(\tau) &= 4\int_0^\infty \frac{\sin^4(\pi f\tau)}{(\pi f\tau)^2}df = \tau^{-1}\,. \end{aligned}$$

The noise sources mentioned so far do not give rise to an Allan deviation with a τ^1 dependence. This power law frequently arises in experimental situations and needs an explanation. The clearest source of a power law of this kind is a ramp in the measured variable; by ramp it is meant that the variable changes linearly with time during the measurement set. This ramp may be due to transient response of the device but in the case of C-II it is found that for the run analysed (TST1-30, 26 July 1997) there was, by chance, a roughly monotonic increase in atmospheric pressure during the $\sim$10 hr measurement time of the data used (see Fig. 4.13).

Noise Type	Source of Noise
White Phase $\sigma_A \propto \tau^{-1}$	Usually due to broadband amplifiers and electronics. Another source is A/D conversion noise due to the discrete nature of the sampling apparatus.
Flicker Phase $\sigma_A \propto \tau^{-1}$	Noisy components cause fluctuations in the electronic properties of components due to thermal effects.
White Frequency $\sigma_A \propto \tau^{-1/2}$	Quantum Noise is the only source for noise obeying this power law. This is due to spontaneous emission in the discharge which produces a random walk in the phase of the lasing mode.
Flicker Frequency $\sigma_A \propto \tau^{0}$	Due to Active Components. Most particularly this noise type has been traced back to the discharge assembly. Bias Instability - flicker in the bias frequency due to unknown non-reciprocal effects.
Random Walk in Frequency $\sigma_A \propto \tau^{1/2}$	Mechanical shocks which are distinct events that knock the apparatus into different operating configurations. Vibrations and micro-seisms are typical sources, usually these will only act as white phase noise sources but may, if large enough, give a continuous supply of mechanical shocks. Temperature and pressure fluctuations also may act in the same manner as mechanical shocks by changing the operating parameters of the laser.
Frequency Ramp $\sigma_A \propto \tau^{1}$	Platform acceleration could cause this but would have to be a very long angular acceleration and apart from tectonic movement this is not considered likely in the ring laser project. The transient response time taken for all parts of the ring to reach equilibrium with the environment may be long enough to give a ramp in the frequency. Essentially any effect of coherence time long relative to the measurement time may cause this; a fraction of a cycle of a very low frequency modulation may appear as a linear trend.

Table 4.1: The different half integer power laws and the typical phenomena that cause them are tabulated here. The power laws are given for the Allan deviation as this is the directly meaningful physical variable.

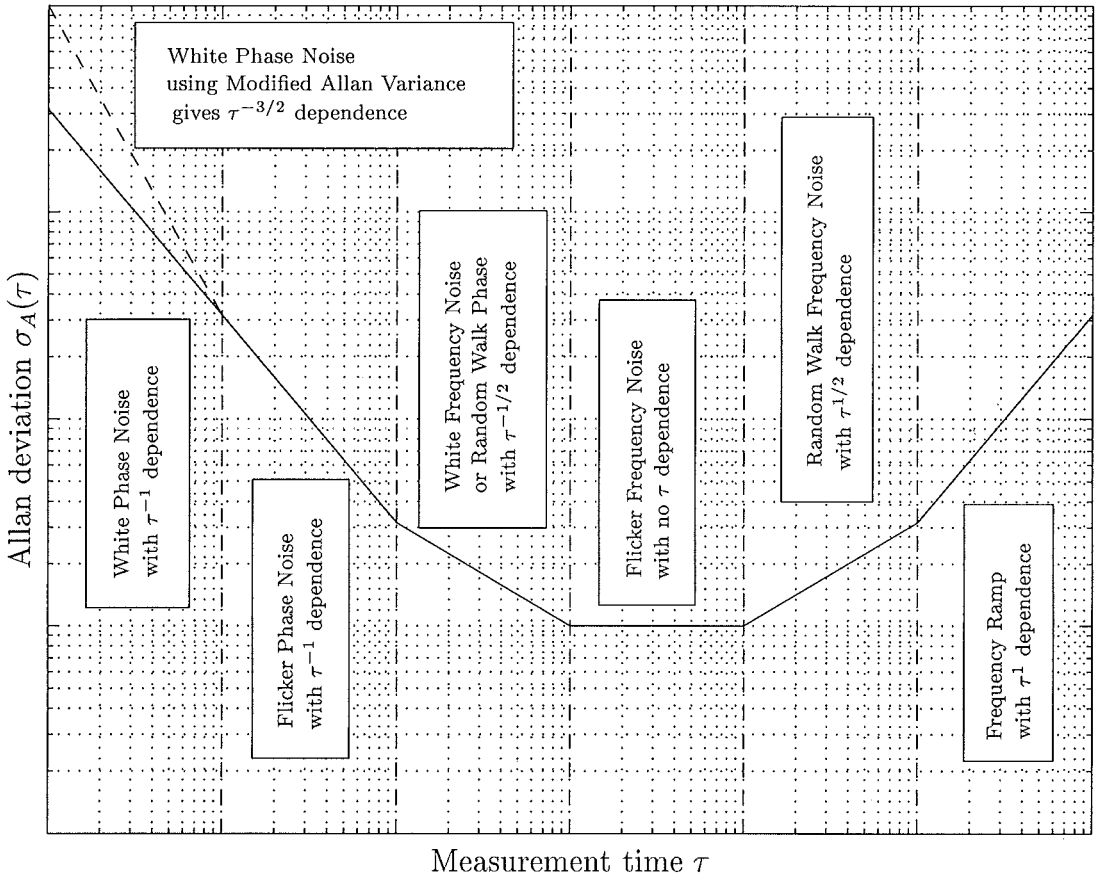


Fig. 4.9: The different noise types give rise to integer or half integer power laws for the Allan deviation as indicated here. Also a linear ramp in frequency appears on this plot to have a linear dependence on τ. The degeneracy between white phase noise and flicker phase noise may be differentiated by use of the Modified Allan variance method.

Correlation between fluctuations in the environmental conditions and the Sagnac frequency are often observed and removal of these dependences is discussed in §4.4.6. A linear drift in the measured variable gives rise to a τ^1 power law. This can be understood by considering the definition of the Allan variance as in Eq. (4.19). In the presence of a frequency ramp the difference in successive measured frequencies increases linearly with the measurement time τ. Hence the Allan variance increases quadratically with τ and the Allan deviation linearly.

4.4.3 Ring Laser Operation at the Quantum Level

The experiment performed by Dorschner et al. [41] demonstrates the operation of a ring laser gyro near the quantum limit. Their ring laser used He-Ne and operated on the 632.8 nm transition. The geometry was triangular with perimeter 55.16 cm. Four mode operation was used with the left and right circular polarisations of a single longitudinal mode oscillating in either direction. The

degeneracy in the optical frequencies was lifted using a passive magneto-optical bias element. Vibration noise sources were made negligible by floating the laser on an air cushion. The output beat frequency was sampled using a frequency meter with mHz resolution and a single 200 s long data record taken. From this record the uncertainty in the measured frequency was obtained for a spread of different sample durations similar to the Allan variance method. Their results (see Fig. 4.10) show that their laser was operated very close to the quantum limit. The authors suggest that the remaining difference is due to a correction of $1 - n_1 g_2 / n_2 g_1$ due to the lower lasing level not being empty; this is discussed in §4.2.2. The theoretical predictions are given by the expression

$$\Delta\omega_{\text{rms}} = \frac{1}{3}\sqrt{\frac{\hbar\omega_0^3}{2Q^2P}}\, T^{-1/2} \tag{4.23}$$

where the factor of $1/3$ comes from the power measurement being taken at only one port and the triangular geometry. The power injected per mode, P, is related to the combined beam output power at each port, P_o, according to $P = 2P_o l/t$. The total losses and mirror transmission are denoted by l and t respectively.

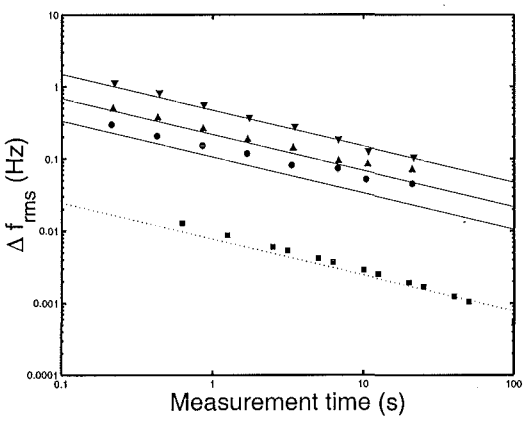


Fig. 4.10: The experimental results from Dorschner et al. [41]. The theoretical lines represent the $T^{-1/2}$ dependence of $\Delta\omega$. The ▼, ▲ and ● represent the Dorschner results for injected power values of 2.7 μW, 13 μW and 52 μW respectviely. The ■ represent C-II data taken from the 26 July 97 data set and the dotted line represents the theoretical quantum limit appropriate for these data.

For comparison the results from C-II are plotted on the same graph which have a *fitted* 0.010 $T^{-1/2}$ power law. These results are obtained by following the same procedure as that taken by Dorschner et al. except that the frequencies are determined digitally using the AR(2) method in post processing. The theoretical quantum limit is indicated with a dotted line. The parameters used to produce this line were taken at different times and by different people and hence I will stop short of saying that the closeness of the line to the data points indicates that the laser was operating at the quantum limit. For a laser of this geometry the power injected per mode is given by $2P_o l/t$ so that

$$\Delta f_{\text{rms}} = \frac{1}{2\pi}\sqrt{\frac{\hbar\omega_0^3 t}{4Q^2P_o l}}\; T^{-1/2}\,. \tag{4.24}$$

By considering the loss per orbit of the cavity the losses may be expressed in terms of the quality factor of the cavity and the free spectral range according to $\omega_0 = 4lQ\nu_{\rm fsr}$, allowing the rearrangement

$$\Delta f_{\rm rms} = \frac{1}{2\pi}\sqrt{\frac{\hbar\omega_0^2 t\nu_{\rm fsr}}{QP_o}}\ T^{-1/2}\ . \tag{4.25}$$

The manufacturer's quoted value for the transmission of the super-mirrors is given as $t = 0.2\,$ppm. The cavity Q is taken to be 5.93×10^{11} which is found from the cavity ring-down measurement of $200\,\mu$s (measurements taken on 17 June 97) and the output power is taken as 60 pW. The free spectral range of the cavity is 75.00 MHz and is known to high accuracy. The expected relationship between the frequency fluctuations and the integration time is thus $\Delta f_{\rm rms} = 0.00774\,T^{-1/2}$ (indicated on Fig. 4.10 by a dotted line). These results are highly consistent with the half integer power law which arises from white frequency noise. §4.4.2 provides a classification list of noise types and there it is seen that the only noise type associated with a $T^{-1/2}$ power law is quantum noise. It is reasonable to conclude that C-II operates at the quantum limit because its Allan deviations follow the $T^{-1/2}$ power law for measurement times shorter than ~100 s.

4.4.4 Using Burg Estimates to Determine Linewidth

The signal from the ring laser is typically a highly coherent distorted sinusoid with additive white noise from the detector and often polluting coherent signals, most prominently from the mains power (50 Hz) and sometimes from other instruments. The distortion gives rise to a complex harmonic structure which is discussed in Chapter 5. Here only the linewidth of the carrier is of interest and hence the data is bandpass filtered removing all spectral features other than the Earth line. In doing this most of the spectrally broad background noise is also removed which provides a much cleaner signal.

The FWHM linewidth, $\Gamma_{\rm Hz}$, is estimated using two entirely different techniques. The first technique assumes that for short integration times (<100 s for C-II) the Allan variance is dominated by the white frequency noise arising from quantum noise. As seen earlier the Allan and standard deviations are equivalent under these circumstances. The formula $\Delta f_{\rm rms} = \sqrt{\Gamma_{\rm Hz}/(2\pi T)}$ provides the power law relating the frequency fluctuations $\Delta f_{\rm rms}$ to the integrating time T. By calculating frequency estimates for every half second interval and clustering 200 such estimates together a linewidth estimate is found for every 100 s segment of data. The second linewidth estimation method models the bandpassed signal with the AR(2) process discussed in §3.6.5 where it is found that Eq. (3.55) gives an estimate of the linewidth only after being adjusted by the procedure described in §3.6.6. Once again 200 estimates are averaged to give a smoother profile and to allow direct comparison with the first method.

Approximately 30 hours of data, taken from C-II starting on 26 July 1997, is

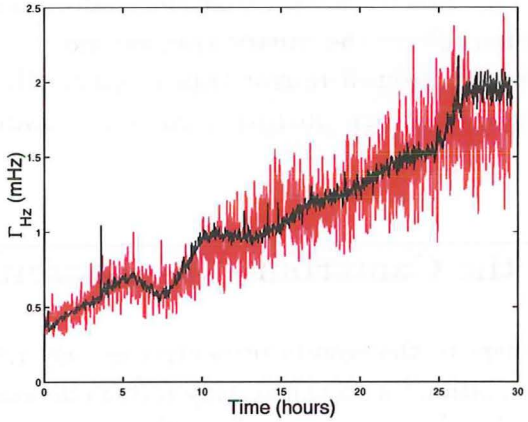


Fig. 4.11: Comparison between two independent methods for estimating the white frequency noise induced linewidth. The results of the AR(2) method are plotted in black while the Allan variance method is shown in red. The Γ_{AR2} estimates are scaled up by a factor of 14.8 to arrive at the Γ_{Hz} estimates. This scaling factor is determined from the noise content and the passband bandwidth B used when reducing the data.

used to compare the two methods using real data. The segments of data used to make estimates are short (1 s) so that there are many estimates of the linewidth all of which are sufficiently short to be effectively independent of any drift that is present. Fig. 4.11 shows the comparison between the two methods. The correction factor which is required to rescale the linewidth estimates is 14.0 (see §3.6.6) It ought to be emphasised that the two methods are entirely independent. The first relies on many frequency estimates which may be acquired via several different techniques, while the second results from choosing the AR(2) process to model the signal. The agreement is quite striking in view of the fact that linewidth measurements were not previously available. The congruence of the inflexions in these two plots suggests that the resolution of either method is sufficient to reveal fine structure.

A primary result that comes from this analysis is that while the Allan variance method requires many frequency estimates before a reasonable linewidth estimate may be found, the AR(2) method provides a linewidth estimate for each 1 s data segment. In the event that only a limited amount of data is available an estimate of linewidth may be made if the signal is of sufficient fidelity. The power of the AR(2) approach is well exemplified by the data analysed here. Given a 1 s data stream (400 samples) of only 80 cycles not only is a $<$ mHz accuracy frequency estimate available but also an estimate of the frequency noise induced linewidth.

In summary, the quantum noise induced linewidth, $\Gamma_{\rm Hz}$, is determined from the experimental data using two methods which agree well. The theoretically expected linewidth is also in good agreement with these measured linewidths. The five quantities $\Gamma_{\rm Hz}, t, Q, P_o, \nu_{\rm fsr}$ are related by the expression

$$\Gamma_{\rm Hz} = \frac{1}{2\pi}\frac{\hbar\omega_0^2 t\nu_{\rm fsr}}{QP_o} \,. \tag{4.26}$$

Three of these $(Q, P_o, \nu_{\rm fsr})$ are readily measured. Previously the manufacturer's quoted value for mirror transmission has been used to evaluate the quantum limits

on performance. No quantitative account has been made for the degradation of mirrors. The added ability to measure Γ_{Hz} directly means that the performance limits are known and the above expression allows the mirror transmission, t, to be estimated indirectly. Monitoring the evolution of mirror transmission allows more informed decisions to be made regarding when mirrors need to be cleaned or replaced.

4.4.5 Comparison Between the Canterbury Ring Lasers

Much of the discussion in this section refers to the results presented in Fig. 4.12. These results depict the relative Allan deviation for the three ring lasers currently (November 1998) in operation at Canterbury and as such allow the comparison between the lasers of their stability as a function of measurement time. No discussion is provided here about the accuracy with which the lasers measure the Earth's rotation rate.

The white frequency noise is shown in §4.4.3 to be due primarily to quantum noise. The root mean square (rms) frequency fluctuations induced by quantum noise obey the relation

$$\Delta f_{\mathrm{rms}} = \frac{1}{2\pi}\left(\frac{\hbar\omega_0^3}{2Q^2P}\right)^{1/2} T^{-1/2} \,. \tag{4.27}$$

The quantum noise content of frequency fluctuations is determined by the operating parameters P, the power injected per mode, and Q, the quality factor of the cavity. Single mode operation of the lasers means that the lasing frequency is never further than half the free spectral range from the atomic line centre, $\nu = 474\,\mathrm{THz}$, and hence all our lasing frequencies are within a factor of 1 ± 10^{-7} of each other. For the purposes of the current discussion the variation of ν between the lasers is negligible. For each of the ring lasers the noise content is dominated by the white frequency noise at short measurement times; this is attributed to the quantum noise described above. Each laser has a different level of quantum noise contribution proportional to Q^{-1} (which is dependent on the mirror quality and cleanliness and the arm length of the laser) and proportional to $P^{-1/2}$.

The quality factor of the cavity is determined using a cavity ring-down measurement where the discharge is turned off rapidly relative to the decay time, by shorting out the rf excitation, and the exponential decay observed. These measurements vary for different alignments of the cavity; very slightly different geometries may utilise different parts of the mirrors and have an effect on the Q measurement. For this reason the Q can only be quoted with large uncertainties as the measurements were not performed simultaneously with the data logging. The ring-down measurements nearest to the time of data collection are summarised in Table 4.2. Observation of C-I and G-0 through the glass lids of the mirror housings clearly shows that a significant fraction of the total mirror

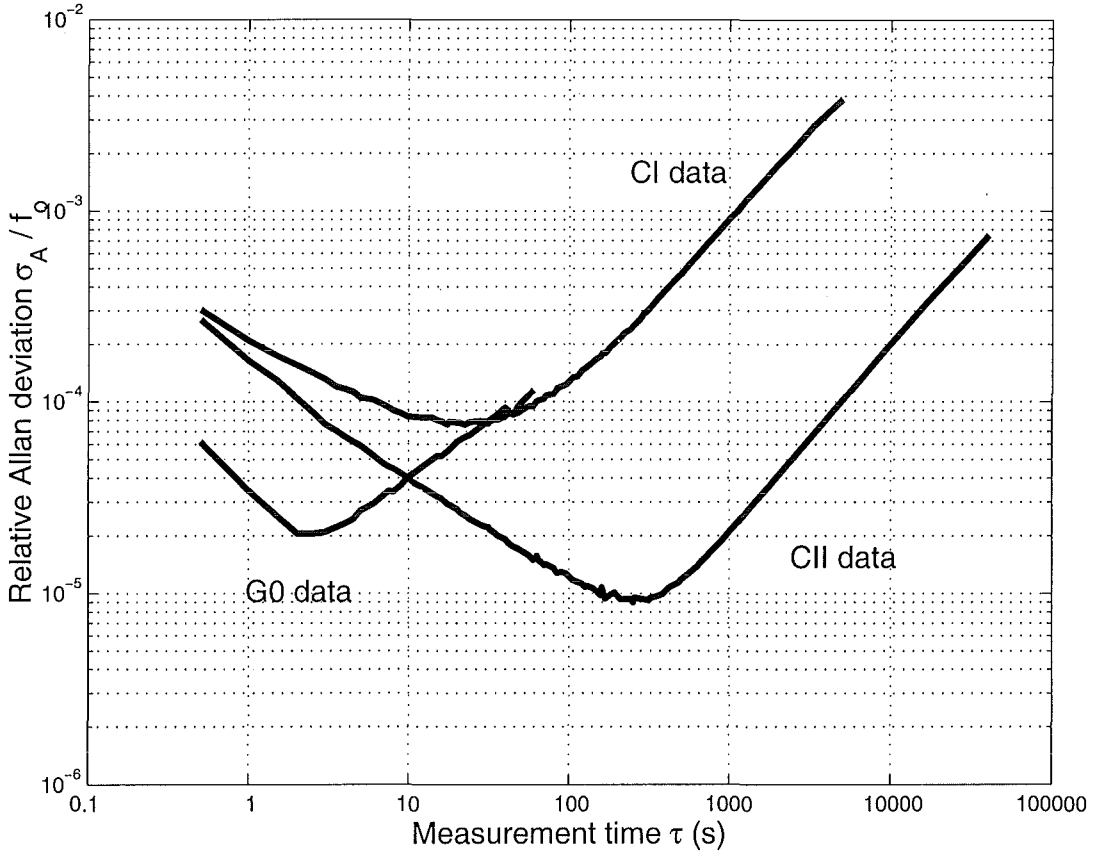


Fig. 4.12: A comparison of the relative performance of the three ring laser gyros operated at Canterbury is provided here. The data for G0 was taken on 29 Jan 1998, for C-I on 12 Oct 1998 and for C-II on 26 July 1997.

losses are due to scatter from the front of the mirror; the front of the mirrors show several bright spots while the exit beam is hardly visible. In order to relate the cavity Q and the output power to the power injected per mode a measurement of the mirror transmission is required. Unfortunately such measurements are not readily performed locally and hence the manufacturer's quoted values for transmission (4 ppm) are adopted.

With these values several qualitative conclusions may be drawn from Fig. 4.12. The quantum noise apparent in C-I and C-II is roughly equal which is expected for two lasers of similar size and with similar mirrors. The environmental noise is approximately 100 times greater in C-I than in C-II which follows as a result of rigidity and temperature insensitivity of the C-II design. G-0 has roughly the same level of environmental dependence as C-I but the quantum noise level is somewhat lower which is attributable to the size increase of the laser. These conclusions illustrate that if stability of 10^{-9} is to be achieved in a reasonable measurement time then both quantum noise levels and environmental dependence have to be addressed. Taking a day ($\approx 10^5$ s) as a reasonable integration time

Laser name	$\tau_{1/e}$ decay time	Cavity Q	P_{out}
C-I	15±5 μs	4.5±1.5×10^{10}	400 pW
C-II	150±50 μs	4.5±1.5×10^{11}	5 pW
G-0	250±70 μs	7.4±2.1×10^{11}	

Table 4.2: Results of ring-down measurements for C-I, C-II and G-0. The $1/e$ decay times are quoted along with the corresponding Q which is assumed to govern the cavity intensity exponential decay according to $I(t) = I_0 \exp(-\omega_0 t/Q)$ so that $Q = \omega_0 \tau_{1/e}$. The last column summarises the typical single mode, single beam output power.

then 10^{-9} stability is only achieved if the quantum noise level is reduced by a factor of 100 from the level observed in G0. Eq. (4.27) indicates that for a laser of given size this may be achieved by increasing Q or P or by decreasing t. The latter is unlikely because the mirrors currently in use are already state-of-the-art. Improved cleanliness of the mirrors will increase Q by reducing cavity losses and is therefore of utmost importance. It is also possible to increase the power injected per mode. One method, used by Dunn [42], increases the gas pressure inside the cavity so that the atomic transition becomes predominantly homogeneously broadened. In this way a single longitudinal mode lases at high power while starving the neighbouring modes of sufficient gain to lase. Another approach is to let the laser oscillate in more than a single longitudinal mode thus allowing the power injected per mode to increase. Multimode operation of the Canterbury Ring Lasers has not been a standard mode of operation and hence very little analysis of multimode behaviour exists.

A few points need to be made regarding the data presented in Fig. 4.12. In order to gain a meaningful comparison between stable oscillators it is important that the correct frequency be used to scale the Allan deviation. It is tempting with a ring laser to divide $\sigma_A(\tau)$ by the optical frequency of the laser, after all, the beat frequency is the frequency difference between the two counterpropogating modes. However, these modes share a common cavity. It is quite reasonable to suppose that substantial fluctuations in the optical frequencies of the modes occur, due to thermal effects say, but that these fluctuations are correlated or in fact equal. When two different Iodine stabilised lasers are combined to give a beat frequency the fluctuation in frequency of the resulting combined beams is of the order of 10 kHz in one second and our lasers have no such stabilisation. See for example the efforts required by Pfleegor and Mandel [91] and Hall [100] to identify beats between lasers which do not share the same cavity. These rapid fluctuations are not seen in the Sagnac signal from any of the Canterbury ring lasers from which it is deduced that they are equal for each mode. If the optical frequencies are indeed fluctuating in unison then our high stability cannot be attributed to optical stability of the modes and $\sigma_A(\tau)$ should not be scaled by the optical frequency. The proper frequency to use as a scaling factor is the Sagnac frequency as this is the frequency from which the fluctuations are found and is approximately 69,79 and 288 Hz for C-I,C-II and G-0 respectively. Output

from these three lasers is illustrated in Fig. 4.12 and here the rescaling uses the audio frequency determined from the data set in question.

One further point needs to be made regarding the graphing method used in Fig. 4.12. The Allan variance is found from a series of N frequency measurements and requires that all the k subgroups have a length that is a divisor of N. In order to maximise the number of plot points the data set is reduced in length to $m < N$ where m is chosen to have many distinct factors. I divide N by the largest factorial q which satisfies $q! < N$; any further factors in $N/q!$ (once rounded down to the nearest integer) are found in the same way and so on. This does not necessarily give the optimum answer but is guaranteed to provide a large number of divisors. For example if 3 000 frequency estimates are available then there are 32 distinct factors including 1 and 3 000 itself. However, dividing by the largest factorial smaller than 3 000, that is 6!, gives 4.1666 and this gives a further factor of 4. Now $4 \times 6!$ is 2 880 which has 42 distinct factors including 1 and itself. Thus by reducing the number of estimates used from 3 000 to 2 880 the number of points on the Allan variance plot is considerably increased.

4.4.6 Removal of Observed Pressure and Temperature Dependence

Data is recorded from C-II in the form of segments of waveform sampled at 400 Hz and typically of duration 20 minutes. The waveforms are saved with the start time, atmospheric pressure, humidity and temperature tagged on to the end. It has been noted that when the average frequency for these segments is compared with the atmospheric pressure there is a strong correlation. The mechanism that provides this connection is undoubtedly the flexure of the window on which the piezo is attached. Fig. 4.13 illustrates the correlations between pressure, temperature and frequency where the equations provided are purely empirical. These provide an independent estimate of the frequency according to the relationship given in Fig. 4.13. Although empirical this estimate may be used to dedrift the output frequency. The coefficients of the polynomial fit change from one data set to another. In the absence of a complete model to describe the effects of pressure changes on the output frequency the above scheme is used to dedrift the data. The resulting Allan variance is shown in Fig. 4.14 with the unaltered Allan variance for comparison. There is a considerable improvement with the minimum of the Allan variance about a factor of 2 below the previous minimum. Initially this minimum occurs at 200 s and has a fractional frequency resolution of 9×10^{-6}. After dedrifting using both pressure and temperature this minimum shifts to 500 s with a fractional frequency resolution of 6×10^{-6}. The success of such methods is monitored by the changing shape of the Allan deviation plot. As one noise source is reduced another will become apparent; the power law that governs this new noise gives information about its source and so this too can be dealt with and so on. Ultimately the performance of the ring will be limited by the white frequency noise arising from spontaneous emission and

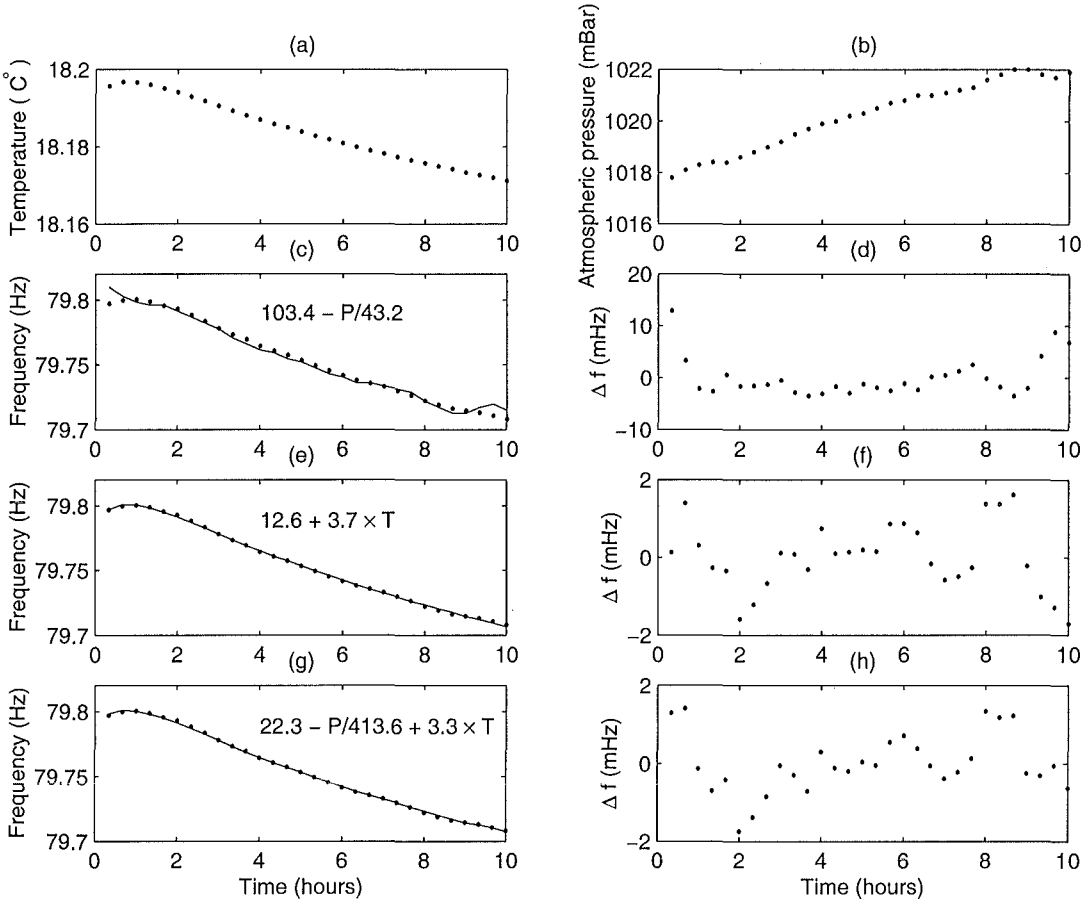


Fig. 4.13: This figure illustrates the dependence of the output frequency on the environmental parameters of pressure and temperature. For the data considered here the correlation between pressure and frequency (c) is not as good as that between temperature and frequency (e). The combined pressure and temperature (g) is not much of an improvement. The residuals of these latter two (f) and (h) are near the resolution of the AR(2) frequency estimation process.

the flicker frequency noise arising from the gain medium although the latter has not been identified in C-II yet [18].

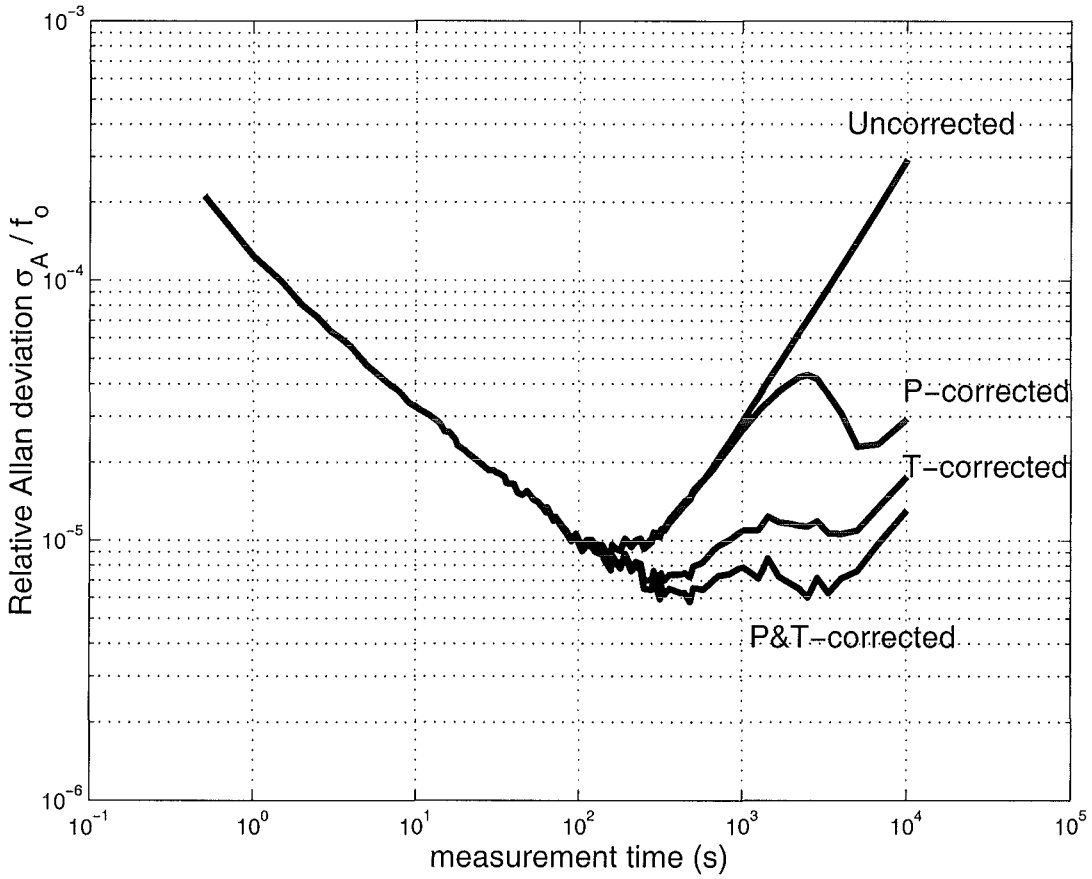


Fig. 4.14: The comparison between the Allan variance of the ring output before and after the pressure correction process. The best improvement occurs for long τ suggesting that the ramp frequency that gives the τ^2 dependence is largely due to pressure effects.

5. Single Beam Output Model

In this chapter a model is developed to describe the oscillations of the single beam irradiances which oscillate at the Earth induced rate. To be successful the model must describe the rich variety of waveforms that are observed in the ring laser output as well as the frequency push/pulling effects. The model is compared with real data as a test of its ability to produce appropriate waveforms with distortion and frequency shifts that are actually observed in real experimental output.

5.1 Introduction

This chapter considers the equations of motion that govern the electric fields and irradiances in ring lasers. These equations, found in § 5.1.6, are constructed in stages with several sections contributing to the final form. In § 5.1.1 field regeneration equations take account of the saturation of the gain medium in a linear manner which is sufficient to provide stable lasing action. A reworking of the regeneration method which includes mode coupling is provided in §5.1.2. The resulting equations are reduced down to the Adler equation and the solution provided. Non reciprocal effects, which arise solely from the geometry of the laser design, are discussed in § 5.1.3. The pivotal effects of backscatter are classified in § 5.1.4 in preparation for its subsequent inclusion into the model. Experimental observations which I have made are documented in § 5.1.5. In § 5.1.6 the backscatter field terms are isolated and inserted into the field equations and three dynamic equations (Eqs. (5.22)) result which are used to describe the ring laser behaviour.

Often when analysing ring lasers multimode dynamics are considered which relate to the case where more than a single mode oscillates in either direction. Statz et al. [107] consider the situation where two pairs of modes oscillate in a single lasing cavity; this is intended as a means of overcoming the lock-in problem in a purely optical manner. Such multimode situations are not discussed here, only the condition where a single mode oscillates in either direction is considered. The two oppositely directed travelling waves in the laser cavity constitute a pair of coupled oscillators. These oscillators may potentially be so strongly coupled that they lock together. This phenomenon was first observed by Huygens [61] who noticed that two pendulum clocks with their own independent swing times may swing in exact unison if mounted close enough together. Understanding of this lock-in phenomenon was extended by Adler [1] who studied the coupling effects on

an oscillating electrical circuit which has a small oscillating signal injected. Adler found that the phase difference between the oscillating circuit and the injected signal is governed by the now famous Adler equation

$$\frac{1}{2\pi}\frac{d\psi}{dt} = a - b\sin\psi \tag{5.1}$$

where a is the frequency difference between the oscillator and the injected signal when decoupled and b is a measure of the coupling between the circuit and the injected signal, later found to be the lock-in threshold. The most noteworthy

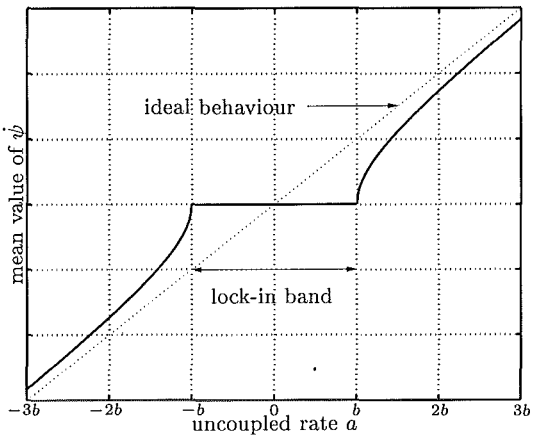


Fig. 5.1: The Adler lock-in curve. It can be seen that the coupling has a considerable effect on the output frequency when near the lock-in region. When $|a| < b$ the oscillators are locked together. When $|a| \gg b$ the output frequency tends asymptotically towards the ideal behaviour which is shown.

feature of the solution of the Adler equation, depicted in Fig. 5.1, is the lock-in band. As observed by Huygens the oscillators lock together when $|a| < b$ and tend asymptotically towards the ideal (no effects of coupling) as $|a| \gg b$. The use of the word *ideal* is subjective and refers to the current context where the ideal gyro is one which has no coupling between oppositely directed travelling waves.

The use of the Adler equation to describe ring laser output is discussed in appendix E. The technique is not successful. If a more complete model is to be found the field equations of the ring laser must be considered more carefully.

In order to investigate the behaviour that arises from coupled laser modes dynamic equations are sought. A satisfactory set of equations may include gain, gain saturation, cross saturation, backscatter, forward scatter and cavity power loss mechanisms. The dynamic equations that govern the fields of the oppositely directed traveling waves may be derived in a variety of ways. Several authors start from Maxwell's equations [6, 9, 126] while other authors start from a set of self consistent equations [7, 8, 10, 11, 28, 45, 57, 90, 101, 112]. The self consistent equations, or regenerative equations, are those that satisfy the optical boundary conditions so that the fast variables return to the same values after each orbit of the cavity. The scalar field amplitudes and differences in phase between oppositely directed travelling waves are slowly varying and are allowed to evolve with time. The latter technique of using regenerative fields is used throughout this chapter because it is easier to work with and is entirely equivalent to the Maxwell equation approach.

Once a set of dynamic equations is found they usually need to be manipulated into a form that lends itself more easily to interpretation and solution. Often solutions are introduced where the fast and slow variables are separated. The intention is to remove the fast variables from the dynamics and this is readily achieved. The solutions that are most commonly used are

$$\tilde{E}_{\pm}(t) = E_{\pm}(t)e^{i(\omega_{\pm}t \mp \phi_{\pm}(t))} \tag{5.2}$$

where the rapidly oscillating complex fields $\tilde{E}_{\pm}(t)$ have the rapidly rotating phasor $\exp(i\omega t)$ factored out isolating the slowly evolving field $E_{\pm}(t)$ and slowly evolving phase $\phi_{\pm}(t)$. This process is adopted by almost all authors and readily gives slowly varying dynamic equations in the scalar field amplitudes and the phase difference between oppositely directed travelling waves ψ where the following definitions are assumed throughout:

$$\psi(t) = \phi_{+}(t) + \phi_{-}(t) + \Omega t \quad \text{where} \quad \Omega = \omega_{-} - \omega_{+} \; . \tag{5.3}$$

Naturally the dynamic equations are best expressed in a variety of forms depending on the assumptions made and processes included in the derivation. Stedman et al. [112] find that their dynamic variables are best chosen as the sum and difference of the field irradiances. These two new variables give simple equations that provide an insight into the dynamics that would otherwise be missed. Wilkinson [126] on the other hand works with the logarithm of the field amplitudes. This choice of variable change is suggested by the frequent appearance of terms such as

$$\frac{1}{E}\frac{dE}{dt} = \frac{d\ln(E)}{dt} \tag{5.4}$$

which will be seen later in this chapter. Although this change of variables provides more simple algebra it is often difficult to keep track of the physical processes behind the calculations. Pesquera and Blanco [90] go further with this theme and define variables solely on the merit of simplifying algebra. None of these schemes is adopted here. Instead the dynamic variables that are used are those that are actually measured namely field irradiances. This choice follows Aronowitz and Collins [7] and to a lesser degree Etrich et al. [45]. Use of the irradiances as the dynamic variables provides a connection with the latter part of the chapter where a comparison is made between the predictions of the model described here and measurements made on C-I.

The solution scheme presented above typically gives rise to three strongly coupled, non-linear, first order differential equations. These equations do not usually lend themselves to closed form solutions, however, under certain conditions it is possible to proceed further. In order to probe deeper into the structure of solutions many authors [45, 57, 112, 126] limit discussion to the conservative (Hermitian) or dissipative (skew-Hermitian) cases. These titles refer to the phase of the backscatter relative to the forward field; conservative backscatter is in quadrature and dissipative backscatter is in phase with the forward field. The origin

of these names is discussed by Wilkinson [126] where coupling between modes is implemented by introducing to the cavity a laminar of material perpendicular to the direction of propagation of the modes. The distinction between the types of coupling is made by the nature of the laminar, dielectric or conductor. To summarise the definitions

conductor - lossy - dissipative - skew Hermitian
dielectric - lossless - conservative - Hermitian .

Having restricted the dynamics to purely dissipative or conservative coupling the solutions may be found in closed form as in Etrich et al. [45].

In general the coupling mechanism is unlikely to be pure; rather a mix of coupling types is more likely the case. The use of dielectric slabs by Wilkinson [126], Haus et al. [57] and Chow et al. [28] is a device to ensure a pure type of coupling. In the case of the dielectric slab the coupling is back reflection as opposed to backscatter and there is no net energy loss from the cavity. The Canterbury Ring Laser uses no optical elements in the ring cavities that could provide such a mechanism; the predominant means of energy loss from the cavity is scatter and transmission through the mirrors. Coupling is thus thought to be mostly due to imperfections in mirrors and particles deposited on the mirror surface. According to Haus et al. [57] such mechanisms that tend to scatter light from the cavity will provide only a relatively small amount of backscatter and are mostly dissipative.

It has been mentioned that the dynamic equations usually reduce to three coupled differential equations (see in particular Aronowitz [7]), two of which govern the oppositely directed travelling wave field amplitudes (or irradiances). The third equation is commonly referred to as the frequency determining equation and governs the evolution of the phase difference, ψ, between the two modes. Often this last equation is decoupled from the field equations by making appropriate assumptions and it is solved separately. For example Stedman et al. [112] arrive at an Adler like equation for both dissipative (taking $E_+ = E_-$) and conservative (taking $|E_+|^2 + |E_-|^2$ as constant) coupling. The solution may then be inserted into the field equations to give a waveform. This decoupling often leads to a frequency determining equation that looks and behaves similar to the Adler equation. The key similarities are the existence of a lock-in band and the possibility only of frequency pulling. Aronowitz et al. [7] derive a quite general frequency determining equation which still gives the Adler expression $\nu_{\text{obs}}^2 = \nu_{\text{ideal}}^2 - \nu_{\text{lock-in}}^2$. This is unsatisfactory as is evidenced by the frequency pushing which appears in all the ring lasers operated at Canterbury. These lasers are capable of producing output frequencies several hertz in excess of the theoretical Earth induced frequency. Aronowitz and Lim [8] discuss the existence of a positive correction to the scale factor of a gyro. To do this they use a model which is a special case of the model developed here. A desirable property of a model therefore is that it does not reduce to an Adler type of frequency determining equation. The model that is deduced and explored here does not lend itself to closed form solutions, only numerical solutions are sought. For this reason no unnecessary assumptions

are made and a richer spectrum of solutions unfolds.

The work contained in this chapter follows the work of Aronowitz and coworkers [6–9, 11] and more recently the work of Etrich et al. [45] and Stedman et al. [112]. The Canterbury ring lasers differ from those studied by Aronowitz in that they are r.f. excited, have exceptionally large quality factors, have ten times the perimeter and are square. These differences need elaboration. The r.f. excitation allows the effects of Langmuir flow to be neglected at least to first order [112, §I]. The increased perimeter gives a smaller free spectral range. Single mode operation is achieved by starving all but one longitudinal mode of sufficient gain to lase so that the small signal gain profile of the plasma needs to be extremely close to the cavity losses at atomic line centre. This, in combination with the high cavity Q, ensures that the steady state gain per pass is much smaller than those considered by Aronowitz.

The main results of Aronowitz that are of interest here are now outlined. In the earlier work [7,10,11] a winking phenomenon is discovered and analysed. Winking refers to the difference in irradiance between oppositely directed travelling waves experienced when a ring gyro is rotated below the lock-in threshold. Once lasing the ring gyro is rotated at different rates ranging from zero to well in excess of lock-in and then back down to zero. It is found that at rotation rates below lock-in the mode that travels with the direction of rotation decreases in irradiance while the other mode increases. The sum of the irradiances remains approximately constant. As the rotation rate increases the co-rotating mode's irradiance decreases monotonically for all discharge currents. Eventually the modes decouple and oscillate with a highly distorted waveform and large irradiance fluctuations (peak to peak). Upon increasing the rate further the irradiance fluctuations decrease and the oscillations tend towards the ideal undistorted sinusoidal profile. When the rotation rate is lowered to the value at which it previously unlocked the gyro continues to oscillate and with much less distortion than before. Eventually the gyro locks again but at a lower rotation rate than that at which it unlocked. This hysteresis is attributed to the gain medium. However, an alternative explanation is provided in §5.3.3 where it is seen that the model described in this chapter also describes the winking phenomenon. The Canterbury ring lasers are normally operated well above the lock-in threshold. Aronowitz also considers the positive scale factor correction [8] which he observes experimentally. This correction is found to increase as the rotation rate is lowered towards the lock-in threshold. At the lock-in threshold the scale factor falls rapidly to zero, the peak scale factor increasing with lower gains. This behaviour is predicted by Etrich et al. [45, Fig. 8] in the conservative coupling regime.

5.1.1 Gain Saturation

In order to introduce the ideas of field regeneration, which are used throughout this chapter, the field equations are deduced here with gain saturation included. The dynamic equations for the electric fields of the lasing modes are determined

by the regeneration approach used by Siegman [104, p.413]. The approach is to determine how the field regenerates itself after a single orbit of the ring cavity and, using the fact that the orbit time is small, take the differential limit. The modes are labelled with a + subscript or a - subscript. In order to avoid confusion between groups working in different hemispheres (or those that have the ability to rotate their lasers in either direction) these subscripts do not refer to a definite sense of rotation but to the modes that travel with (+) or against (-) the direction of rotation of the laser. For simplicity the saturation of the gain is taken to be homogeneous and hence, using the saturation expression from Siegman [104, p.207], the field regeneration equations are

$$\tilde{E}_{\pm}(t) = \left[\tilde{E}_{\pm}(t-\tau) + \frac{\tau}{2}\left(\frac{\alpha_{0\pm}}{1+I_{\pm}(t)/I_{\mathrm{sat}\pm}} - l_{\pm}\right)\tilde{E}_{\pm}(t-\tau)\right] e^{i\omega_{\pm}\tau} \,. \tag{5.5}$$

where the cavity gain has been written in terms of the small signal power gain, $\alpha_{0\pm}$, and the saturation irradiance, $I_{\mathrm{sat}\pm}$. These terms require some explanation. On the left of the equation is the field at time t. On the right are all the field components that combine to give rise to $\tilde{E}_{\pm}(t)$. These are the field components from the previous orbit, at time $t-\tau$, which will undergo loss and gain in the upcoming orbit. The fields must be advanced in phase by $\exp(i\omega_{\pm}\tau)$ to account for the propagation around the cavity. Equating these two forms is a statement of mode regeneration; the lasing boundary conditions are satisfied. The cavity loss is represented by $l_{\pm}$. The loss of energy due to scattered light is dealt with later and hence the losses referred to here are transmission and absorption at the mirrors. Both the loss and gain are coherent with respect to the cavity mode and hence are in phase with the primary phasor $\tilde{E}_{\pm}(t-\tau)e^{i\omega_{\pm}\tau}$. Also these terms include a τ factor as the constants $l_{\pm}$ and $\alpha_{\pm}$ refer to the loss and gain per unit time. The typical manipulations are now worked through and the field equations are derived.

In a ring laser the orbit time τ is sufficiently small that the second term on the right hand side of Eq. (5.5) is small compared with the first. Hence the use of the approximation $\exp(x) = 1 + x$ allows the simplification

$$\tilde{E}_{\pm}(t) = \tilde{E}_{\pm}(t-\tau)\exp\left(i\omega_{\pm}\tau + \frac{\tau}{2}\left(\frac{\alpha_{0\pm}}{1+I_{\pm}(t)/I_{\mathrm{sat}\pm}} - l_{\pm}\right)\right)$$

and separating the fast and slow variables in the manner shown in Eq. (5.2) this equation reduces to

$$\begin{aligned} E_{\pm}(t)\exp\left(i\omega_{\pm}t \mp i\phi_{\pm}(t)\right) &= E_{\pm}(t-\tau)\exp\left(i\omega_{\pm}(t-\tau) + i\omega_{\pm}\tau \mp i\phi_{\pm}(t-\tau)\right. \\ &\quad \left. + \frac{\tau}{2}\left(\frac{\alpha_{0\pm}}{1+I_{\pm}(t)/I_{\mathrm{sat}\pm}} - l_{\pm}\right)\right) \\ E_{\pm}(t) &= E_{\pm}(t-\tau)\exp\left(\mp i\phi_{\pm}(t-\tau) \pm i\phi_{\pm}(t)\right. \\ &\quad \left. + \frac{\tau}{2}\left(\frac{\alpha_{0\pm}}{1+I_{\pm}(t)/I_{\mathrm{sat}\pm}} - l_{\pm}\right)\right) . \end{aligned}$$

With a few trivial manipulations and taking real and imaginary parts the slowly varying field amplitudes and phases are found to obey the relations

$$\begin{aligned}\frac{E_\pm(t)-E_\pm(t-\tau)}{\tau} &= \frac{1}{2}\left(\frac{\alpha_{0\pm}}{1+I_\pm(t)/I_{\rm sat}}-l_\pm\right)\\ \frac{\phi_\pm(t)-\phi_\pm(t-\tau)}{\tau} &= 0\,.\end{aligned}$$

Clearly these expressions are candidates for application of the differential limit. However, τ has a definite value for each ring laser, in the case of the Canterbury ring lasers this is of the order of 10^{-8} s. Thus the derivative that occurs in the differential equation that results should be understood only to apply to processes that change slowly relative to this time scale. With this reservation the differential limit is taken and is denoted by ∂_t to distinguish it from real differentiation. The trivial solution for the phase evolution is found as $\partial_t\phi_\pm = 0$ so that the phase difference between the two modes evolves linearly with time (remembering the definitions of Eq.(5.3)) and the modes are completely decoupled. This comes as no surprise as no mechanism has been included yet to couple the two modes together. The time evolution of the field irradiance is then

$$\partial_t I_\pm = \left(\frac{\alpha_{0\pm}}{1+I_\pm(t)/I_{\rm sat}}-l_\pm\right) I_\pm$$

where $(\partial_t I_\pm)/I_\pm = 2(\partial_t E_\pm)/E_\pm$. This differential equation could have been stated at the beginning of this section as it merely says that the rate of increase of the cavity field irradiance is equal to the excess of gain over loss. However, it has been shown *en route* how Siegman's regeneration equations are used to derive the dynamic equations of the cavity irradiances.

5.1.2 Derivation and Solution of the Adler Equation

Many descriptions of mode coupling in ring lasers result in an Adler type of frequency determining equation. These derivations are mostly congruent and a close inspection is provided here. This provides a background to understanding the behaviour of the dynamics of the ring laser. Later in § 5.1.5 experimental observations are compiled which suggest corrections to the simple model. Thus what follows may be thought of as a first order model.

As a starting point Eq. (5.5) is adopted with the coupling term between modes added. This gives the fast field equations

$$\begin{aligned}\tilde{E}_\pm(t) &= \tilde{E}_\pm(t-\tau)e^{i\omega_\pm\tau}+\frac{\tau}{2}\left(\frac{\alpha_{0\pm}}{1+I_\pm(t)/I_{\rm sat}}-l_\pm\right)\tilde{E}_\pm(t-\tau)e^{i\omega_\pm t}\\ &\quad +r_\mp\tau E_\mp(t-\tau)e^{i\epsilon_\mp}\end{aligned}$$

where $r_\pm$ is the fraction of the field scattered backwards and $\epsilon_\pm$ is the phase difference between the forward field and backscattered field. The separation of

fast and slow variables, given in Eq. (5.2), is again used and the differential limit is taken. The manipulations performed require some care and are followed through in detail. The term $\alpha_{0\pm}/(1+I_\pm(t)/I_{\text{sat}\pm})+l_\pm$ is written as $\beta_\pm$ for brevity. It may be noted that there is time dependence in $\beta_\pm$ due to the presence of $I_\pm(t)$. It is not necessary to follow this time dependence through the calculations as at no time do two terms of the form $I_\pm(t)$ and $I_\pm(t-\tau)$ combine to form a derivative. In the following manipulations it will be seen that the variables $E_\pm(t)$ and $\phi_\pm(t)$ both form derivatives and hence their time dependence is followed throughout. The presence of τ in the saturation term allows it to be carried into the exponent so that

$$\tilde{E}_\pm(t) = \tilde{E}_\pm(t-\tau)e^{i\omega_\pm\tau+\tau\beta_\pm} + r_\mp\tau\tilde{E}_\mp(t-\tau)e^{i\epsilon_\mp} \tag{5.6}$$

The separation of the fast and slow variables allows the fast variables to be removed according to the following steps

$$\begin{aligned}
E_\pm(t)e^{i(\omega_\pm t\mp\phi_\pm(t))} &= E_\pm(t-\tau)e^{i\omega_\pm\tau+i\omega_\pm(t-\tau)\mp i\phi_\pm(t-\tau)+\tau\beta_\pm} \\
&\quad +r_\pm\tau E_\mp(t-\tau)e^{i\omega_\mp t\pm i\phi_\mp(t-\tau)-i\omega_\mp\tau+i\epsilon_\mp} \\
E_\pm(t) &= E_\pm(t-\tau)e^{i(\mp\phi_\pm(t-\tau)\pm\phi_\pm(t))+\tau\beta_\pm} \\
&\quad +r_\mp\tau E_\mp(t-\tau)e^{i(\omega_\mp t-\omega_\pm t\pm\phi_\mp(t-\tau)\pm\phi_\pm(t)-\omega_\mp\tau+\epsilon_\mp)} \\
E_\pm(t) &= E_\pm(t-\tau)e^{(\mp i\partial_t\phi_\pm+\beta_\pm)\tau} + r_\mp\tau E_\mp(t-\tau)e^{i(\pm\psi+\epsilon_\mp)}
\end{aligned} \tag{5.7}$$

where the last step has made use of the definitions in Eq. (5.3) and the term $\omega_\mp\tau$ has been neglected as it deviates only very slightly from an integral number of full cycles. There are two exponentials in this last expression which are treated differently. The first one on the left is expanded according to $e^x = 1+x$ as the argument is small by virtue of the τ. The second exponential is not small but is pure imaginary allowing the trigonometric expansion. Taking real and imaginary parts and forming the appropriate derivatives yields the dynamical equations:

$$\partial_t E_\pm = \frac{1}{2}\left(\frac{\alpha_{0\pm}}{1+I_\pm(t)/I_{\text{sat}}} - l_\pm\right)E_\pm + r_\mp E_\mp\cos(\pm\psi+\epsilon_\mp) \tag{5.8}$$

and

$$\partial_t\psi = \Omega - (\rho_+\sin(\psi-\epsilon_+) + \rho_-\sin(\psi+\epsilon_-)) \quad \text{where} \quad \rho_\pm = r_\pm\frac{E_\pm}{E_\mp}\,. \tag{5.9}$$

This last equation is identical to that of Aronowitz and Collins [11, Eq. (7)] and is very similar to the Adler equation form but requires certain assumptions before the similarity becomes complete. The first of these assumptions requires the fields to have equal amplitudes for all times, $E_+ = E_-$. Also the complex fractional scatter coefficients between modes are required to have the same magnitude, $r_+ = r_-$, and opposite phase, $\epsilon_+ = -\epsilon_-$. The last two conditions expressed here are neatly satisfied if the backscatter is provided by a dielectric slab inserted into the cavity. This constitutes conservative coupling where all energy lost from one mode is gained by the other [57].

For completeness the solution to the Adler equation is included here. There are two solution regions corresponding to the oppositely directed travelling waves being locked or unlocked which are looked at separately. First the case where the oscillators are locked, $|a| < b$, is considered. In this region it is found that $\partial_t \psi = 0$ so that the solution is $\psi = \sin^{-1}(a/b)$. Note that although the oscillators are locked there is still a shift in phase between the two modes which is related to the rotation rate. In this way a locked active ring laser is still rotation sensitive in a passive sense. In the unlocked region $|a| \geq b$ and the Adler equation is rewritten as an integral equation

$$t = \int \frac{d\psi}{a - b\sin\psi} \tag{5.10}$$

$$= \frac{2}{\sqrt{a^2 - b^2}} \tan^{-1}\left(\frac{a\tan\psi/2 - b}{\sqrt{a^2 - b^2}}\right) . \tag{5.11}$$

It is useful to define a new parameter according to $c = \sqrt{a^2 - b^2}$ so that

$$\psi = 2\tan^{-1}\left(\frac{b + c\tan(ct/2)}{a}\right) . \tag{5.12}$$

This solution governs the evolution of the phase between the two modes and may be used to find the pulled (reduced) frequency difference between the two modes. If the pulled frequency has period τ then $\psi(t = 0) = \psi(t = \tau) - 2\pi$ which leads to $\tan(c\tau/2) = 0$ and hence $\pi = c\tau/2$ so that the parameter introduced above, c, is the mean pulled angular frequency. The pulled frequency is presented in Fig. 5.1 as a function of the uncoupled frequency. It should be pointed out that the output frequency can never be greater than the decoupled *ideal* frequency. Knowledge of the exact waveform expected due to coupling can be used to correct a gyros' measurement of rotation rate. This idea is central to the analysis presented by Stedman et al. [112] where good agreement between theory and experiment is reported for data taken from C-I. The analysis is repeated in Appendix E for data sets taken from both C-I and C-II. From this analysis it becomes clear that the Adler equation provides an incomplete description of the ring laser dynamics.

The Adler description is attractive due to its simplicity but is unable to describe the real system. In order to arrive at the Adler form several assumptions are made which are clearly incorrect. Observations of our ring lasers indicate which assumptions cannot reasonably be made. The increased complexity of the model makes interpretation difficult but is necessary in order to improve its performance.

5.1.3 The Inequality of Irradiances in Oppositely Directed Travelling Waves

Often the assumption is made that the counter-propagating modes have equal irradiances. This assumption is discussed by Statz et al. [107, § 4.3 et seq.] where

non-reciprocal gain saturation is found to arise from the introduction of an aperture. Such an aperture which significantly cuts a mode's waist produces diffraction effects (forward scatter) that contribute to its own lasing action. This is quite distinct from backscatter. As discussed by Statz et al. when an aperture (knife edge) is introduced to one side of the gain medium a differing amount of diffraction occurs for the oppositely directed travelling waves. This affects the gain saturation and ultimately the irradiance of the modes. Further, the amount of dispersion experienced by each mode also differs and frequency splitting occurs. A summary of the main points from Statz et al. is given

- The aperture effect vanishes for symmetric disposition of aperture(s) and gain medium.
- The aperture effect exhibits oscillatory dependences on longitudinal and transverse displacements which smooth considerably for a longer gain region.
- The aperture effect depends on the resonator configuration.
- The aperture effect is approximately proportional to the diffraction loss introduced into the fundamental mode for the initial *small-loss* regime.

Furthermore, the effect of gain medium lensing coupled with the aperture effect can give interesting results. Essentially the change in gain with distance from the optical axis gives a changing refractive index which may act as a lens. In a series of experiments by Garside [63] a single isotope ring laser was used with a variable aperture and the change in irradiance of oppositely travelling waves moved from the extreme with unidirectional lasing to bidirectional lasing and then to unidirectional lasing in the other direction. This happens as the laser is tuned from below atomic line centre to above. Almost pure unidirectional lasing has been observed in both C-I and C-II.

The configuration of C-I used for the current analysis is $\begin{smallmatrix} \mathrm{F} & \mathrm{F} \\ \mathrm{C} & \mathrm{C} \end{smallmatrix}$. This provides high symmetry with regard to diffraction losses in the gain constriction. This is the most important location as this represents the smallest aperture the beams pass through. If the beams are slightly misaligned when passing through the gain tube they will see different apertures upon entering and exiting the constriction. An analysis of the extent to which these effects are significant has not yet been carried out. For the purposes of the model described in this chapter the non-reciprocal effects introduced by this asymmetric geometry are neglected. I have chosen a gain tube with sufficiently large internal diameter (7 mm) and ensured that the alignment is axial so that very little diffraction takes place. For these assumptions to hold the alignment of the mirrors is critical and it is essential that the cavity lasing resonance uses the proper parts of the mirrors, namely the centre, so that non-reciprocal diffraction effects can be ignored. It has been observed occasionally that the spots on the mirrors, which are seen as the result of scattered light, are not located in the centre of the mirror and a complete

mirror realignment is required. Once the alignment requirements are satisfied the saturation irradiances may be taken as equal for the oppositely directed travelling waves and $I_{\text{sat}+} = I_{\text{sat}-} = I_{\text{sat}}$.

5.1.4 Sources of Backscatter

The model presented in this chapter concentrates on the coupling between modes which is due to scattered light. Although the model itself uses a somewhat simple approach a collection of the different types of scattering is provided here to help to justify the assumptions made later. A survey of scattering mechanisms is provided by van de Hulst [118]. Fig. 5.2 illustrates the different types of scatter that may

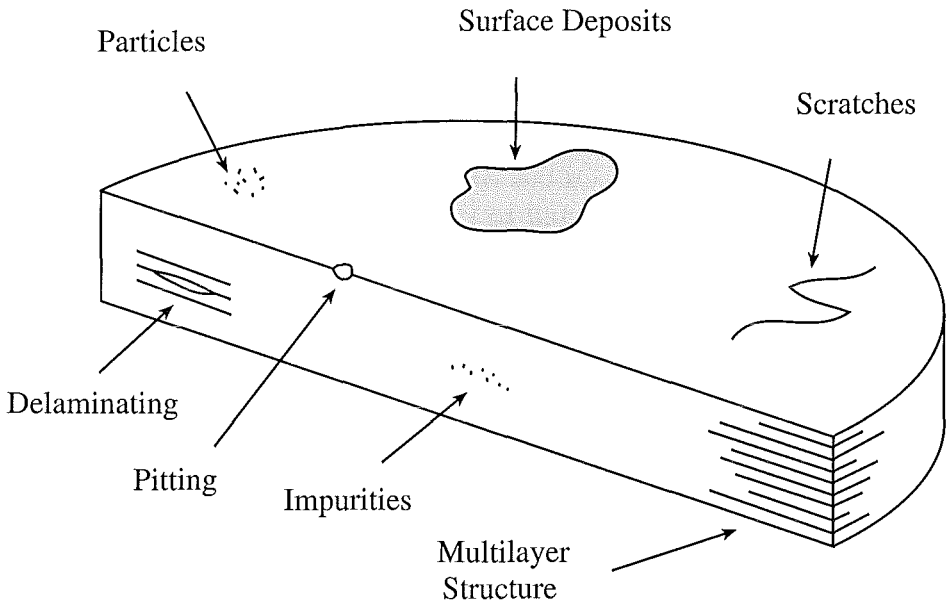


Fig. 5.2: Cross section of the multilayer coating that forms a super mirror. This illustrates the various types of scatter that may be present in the ring laser. The multilayer stack is mounted on a blank usually made of Zerodur™. The thickness of the coating relative to the diameter is greatly exaggerated.

occur. Scatter is visible on the mirrors of C-I while the ring is lasing by looking through the glass lids of the mirror housings. It is important to note that far more light is lost from the cavity due to scatter than due to transmission through the mirror. Access to the scattered light is not available with C-II although the scatter centres on C-II are visible from behind the mirror. The most apparent source of scatter is due to particles attached to the surface of the mirror. Two different types of particle matter are discernible, fibrous material which is easily removed and crystalline *rocks* that are firmly attached to the mirror. These *rocks* are held firmly in place by electrostatic forces. They may arise from impurities in the gas supplies or dust that is long term resident in the tubes and housings that form the vacuum manifold. The greatest likelihood is that dust settles on the mirror while

in transit to the mirror housing. It has been observed that a mirror that has only one or two visible surface particles while in the mirror cleaning suite may have several by the time lasing is started; although this change is due in part to the difference in incident radiation between the cleaning suite and the ring laser itself. The clean room facilities minimise this contamination but can never completely prevent it. Other surface deposits are also known to affect the performance of mirrors. Yamada et al. [127] discuss the deposition of Carbon atoms on the surface of super mirrors in a free electron laser. The deposit is in the form of a film over the surface of the mirror which increases absorption and affects the optical nature of the multilayer stack. These authors achieve excellent restoration of the mirror by an oxygen plasma treatment. The mirror is immersed in an r.f. excited plasma which successfully etches the carbon layer away without seriously damaging the original mirror surface. Such deposits are observed in C-I on the inside of the gain tube and appear as a smear of black-brown. It is supposed that these smears are due to the cracking of long chain hydrocarbons in the r.f. plasma. A possible source is the Viton O-rings used to seal the vacuum chamber. After pumping for some weeks at ultra-high vacuum and with the r.f. field still applied the deposits fade and are eventually undetectable. However, there is no reason to think that these deposits are entirely restricted to the region around the gain tube. Thus a film may be formed on the surface of a mirror during the operation of the ring laser. Gradual reduction of the performance of the ring laser, in particular due to the mirrors becoming contaminated, has been a long standing theme and has required that the mirrors be cleaned often. The cleaning process discussed in Chapter 2 involves a sequence of washes using different solvents and detergents which have been successful in restoring the performance of the ring. Presumably any surface film is removed during this process. However the cleaning procedure involves scrubbing the surface with acetone and a cotton bud. Any particles present during this stage may cause abrasion and give rise to another type of non-specular reflection in the form of scratches. There is no known treatment for removing scratches on a super mirror but such blemishes have not been apparent to any significant degree on any of the C-I mirrors. This is not surprising when it is realized how hard TiO_2 and SiO_2 are. On Mohs' hardness scale (Talc=1 ... Diamond=10) TiO_2 (Anatase) has hardness 5.5-6 while SiO_2 (Quartz) is 7 [124]. Another type of damage that involves irreversible removal of part of the coating is the appearance of pitting. This results from sharp impacts on the mirror. The possibility of such impacts arises if the ultra-high vacuum (10^{-7} mBar) or even low vacuum (3 mBar) is suddenly violated and vented to the atmosphere, incoming particles will be accelerated up to high speed directly towards the mirror surface. There have been a number of occasions where this may have happened in C-I and pitting may have occured on one set of mirrors. Under extreme circumstances it is possible for the alternate layers of TiO_2 and SiO_2 to delaminate. This represents irreversible damage to that part of the mirror and has not been seen at all with any of our super-mirrors.

The fundamental limit of the reflectivity of a super-mirror is discussed by Bilger et al. [20]. These authors show that if the above effects are able to be reduced

sufficiently then the limits are determined by thermodynamic density fluctuations. They find that for the materials most commonly used (TiO_2 and SiO_2) the fundamental limit of $1 - R$ (R is the reflectivity) is $\sim 10^{-9}$. The current state-of-the-art mirrors are 3 orders short of this mark due to the technological problem of reducing impurities in the materials of manufacture.

5.1.5 Experimental Observations of C-I Output

While operating C-I several features of its behaviour have provided impetus for a new model for the dynamics of the field irradiances. The three main observations are, broadly speaking, problems with beam combination, the phenomenon of frequency pushing and the observed modulation of the scattered light at the Earth rotation induced beat rate. The first of these observations is that of the process of beam combination. This is achieved in C-I by using a 50% beam splitter mounted on an assembly which I have designed. This assembly is fully orientable and comprises a kinematic table mounted on a spring loaded turntable which is adjusted using a vernier calibrated screw. Thus good alignment of the two beams is achieved[1]. The model that is derived in the next section does not require the combination of the beams at all. The model describes the independent irradiance fluctuations of the separate beams, which oscillate at the Earth rotation induced beat rate. This lack of an alignment procedure requires more simple apparatus, which is robust and less temperature dependent, and removes several alignment parameters from the model.

The second observation is the frequency pushing which is very often apparent in both the ring lasers. In his thesis Ziyuan Li [128, Chapter 8] discusses the means by which dispersion in the gain medium gives rise to pushing of the Earth rate. Such effects are shown not to provide a large enough pushing effect to explain the observed beat frequency. Several authors [9, 45, 112] discuss the mechanism of mode pushing as a result of backscatter. Etrich et al. in particular provide a thorough analysis. This analysis gives rise to pushing effects when the coupling between oppositely directed travelling waves is conservative. However this is unsatisfactory as the pushing found is not sufficient to describe the C-I output. Also the assumption of conservative coupling is not realistic in either of the ring lasers (especially C-I) as scattered light from pits and particles [57] is a primary loss mechanism. This suggests that the ring lasers are more likely in the non-

[1] For the alignment to be essentially perfect it is sufficient only to have a single fringe, formed by the two beams, across the surface of the detector. Thus an alignment angle which is better than θ where $d\sin(\theta/2) = \lambda$ is required. Taking d as the diameter of the window of the Hamamatsu PMT (~5 mm) gives $\theta \approx 2.5 \times 10^{-4}$ rad. The alignment of the C-I beams is achieved by overlapping beam spots first at a location near to the beam combiner and secondly at a distance of about 2 m. With both aligned the uncertainty in the overlap of the spot centres is ± 0.1 mm and ± 0.3 mm at the near and far locations respectively. This gives an angular alignment uncertainty of $\pm 2 \times 10^{-4}$ rad which satisfies the criterion that the alignment be better than θ. The alignment achieved by the beam combination system used by C-II is somewhat better than this.

conservative coupling regime.

The final observation that suggests a new dynamical model is the modulation of scattered light. It is possible to observe the Earth rotation induced frequency in the scattered light by monitoring it with a PMT mounted above the glass lid of a mirror housing. The Earth rate observed in this manner is the same as that observed in the beams that exit through the back of the corner mirrors. The individual scatterers are clearly visible on the surface of the super mirrors even after the mirrors have been thoroughly cleaned. This modulation may be attributable to the irradiance fluctuations of the individual beams but interferometric causes cannot be ignored. This is especially true for scatterers which are smaller than a wavelength and hence scatter light at a rate dependent on the passage of bright and dark fringes. Wilkinson [126] mentions the indirect coupling between oppositely directed travelling waves which is due not to light being scattered from one mode into the other but rather to light being scattered laterally from the cavity, as observed, at a rate that is dependent on the phase difference between the two modes, ψ. All of the above considerations are included in the model that follows in the next section.

5.1.6 Backscatter Model of the Ring Laser

Here an analysis is provided which is closely related to the analysis given by Etrich et al. [45] but which takes the phenomena discussed in §5.1.5 into consideration. The scatter terms are found initially which are then inserted into the field equations. There are two separate scatter terms that have to be found, the backscatter which directly couples the two modes and the scatter loss from the cavity. These expressions are inserted into the regenerative field equations which are then manipulated into the dynamic equations for mode irradiances (Eq. (5.22)).

The fast oscillating fields, written with a tilde above them, are shown in Fig. 5.3. Notice that allowance is made for both backscatter and forward scatter (see Siegman [104] p.705 for a discussion of forward scatter). This allows the irradiance of light scattered out of the cavity to modulate at the Sagnac rate as is observed. The quality factor Q of the cavity is thus dependent on the location of nodes with respect to individual scatterers; the Q also oscillates. In exactly the same manner the scattered light associated with mode coupling, backscatter, is modulated at the Sagnac rate. The scatter field term for many scatterers is

$$\tilde{E}(\theta,\phi) = \sum_j \left(\boldsymbol{\delta}_{j+}(\theta,\phi)\tilde{E}_+ + \boldsymbol{\delta}_{j-}(\theta,\phi)\tilde{E}_- \right) \tag{5.13}$$

where the phase ϵ_j is contained in the complex $\boldsymbol{\delta}_j$ coefficients so that $\boldsymbol{\delta}_j = \delta_j \exp(i\epsilon_j)$. It is assumed that there are enough scatterers of varying scattering profiles that the two sums $\sum_j \boldsymbol{\delta}_{j+}(\theta,\phi)$ and $\sum_j \boldsymbol{\delta}_{j-}(\theta,\phi)$ may be assumed to

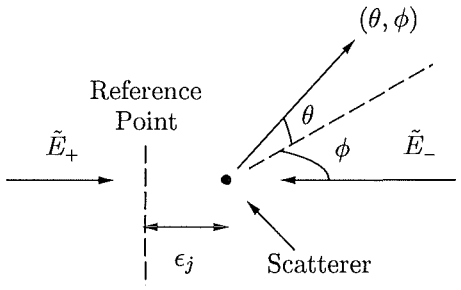


Fig. 5.3: The reference point moves with the body of the ring and is taken as the exit point of the counter-propagating beams. The phase ϵ_j of the scatterer is $2\pi d/\lambda$ where d is the distance of the scatterer from the reference point along the axis of beam propagation.

be uncorrelated. This allows the scatter amplitude to be expressed

$$\tilde{E}(\theta, \phi) = \tilde{E}_+ \delta_+(\theta, \phi) + \tilde{E}_- \delta_-(\theta, \phi) \; . \tag{5.14}$$

There are two different scatter terms that are required. The first is the backscatter which directly couples the two modes together. The second is the coherent scatter loss from the cavity; this must be considered separately from the transmitted loss through mirrors as it is subject to fluctuations at the Sagnac rate. The general equation for the scatter field given above makes the first of these simple to identify, the field scattered into each mode is

$$\begin{aligned} \tilde{E}(d\Omega_\pm) &= \tilde{E}_+ \delta_+(d\Omega_\pm) + \tilde{E}_- \delta_-(d\Omega_\pm) \\ &= \delta_{+\pm} \tilde{E}_+ + \delta_{-\pm} \tilde{E}_- \\ &= \delta_{\pm\pm} \tilde{E}_\pm + \delta_{\mp\pm} \tilde{E}_\mp \end{aligned} \tag{5.15}$$

where the subscripts on the δ coefficients signify first the mode from which light is being scattered and second the field into which the light is scattered.

The scatter loss from the cavity modes is handled differently. In this case energy is coherently lost from the cavity so that a decay term is required. The decay term will have to be dependent on the total energy lost in the scatter field by each mode. The rate of energy loss is found by integrating the irradiance over a surface containing all the scatterers but excluding the solid angles $d\Omega_\pm$ which relate to backscatter. However, it is a non-trivial matter to ascribe a particular fraction of the energy loss to a particular mode. The question that needs to be answered is more clearly understood by examining Fig. 5.4. This experimental arrangement is that used by Pfleegor and Mandel [91,92] to demonstrate the interference of light from two different light sources. Here the lasers are assumed to be sufficiently frequency stabilised that a fringe pattern is clearly visible, this was not the case for Pfleegor and Mandel and they had to resort to extremely short snapshots of the image amounting to only a few photons. The fringe pattern expected on the screen may have the spacing between fringes varied by altering θ. A detector that is placed at the screen may detect the fringe pattern if the fringes are sufficiently large or the detector is sufficiently small. Having achieved this the problem is to determine what fraction of the energy detected comes from each laser. The expectation value of n, the number of photons detected, is shown by Pfleegor and Mandel to be proportional to $I_1 + I_2$ where I_1 and I_2 refer to the irradiances at

the detector when the other laser is occulted. Thus the fraction of the energy detected which has come from each laser is $I_{(1,2)}/(I_1 + I_2)$ where $(1, 2)$ denotes the laser in question. The irradiance of the scattered field is required. Insertion

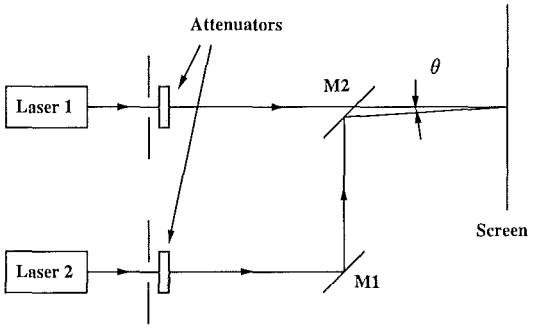


Fig. 5.4: Experimental arrangement used by Pfleegor and Mandel to demonstrate the interference of light from two different optical sources. The beams are misaligned by an angle θ in order to provide a fringe pattern on the screen.

of the solutions from Eqs. (5.2) into Eq. (5.14) and forming the irradiance gives

$$I(\theta,\phi) = \delta_+^2(\theta,\phi)I_+ + \delta_-{}^2(\theta,\phi)I_- + 2\sqrt{I_+I_-}\delta_+(\theta,\phi)\delta_-(\theta,\phi)\cos(\psi + \epsilon(\theta,\phi)) \tag{5.16}$$

where $\epsilon = \epsilon_+ - \epsilon_-$. Thus the irradiance lost from either mode is given as the appropriate fraction of this irradiance. The total energy lost from each mode due to lateral scatter is then found by integrating over a surface that contains all the scatterers. The total energy lost from either mode between the times $t - \tau$ and t is thus

$$\tau \oint \frac{I_\pm \delta_\pm^2(\theta,\phi)}{I_+\delta_+^2(\theta,\phi) + I_-\delta_-^2(\theta,\phi)} I(\theta,\phi)\frac{d\Omega}{4\pi} . \tag{5.17}$$

This energy loss is to be compared with the energy lost from a decaying mode with decay constant γ. The regeneration equations for such a simple case are

$$\tilde{E}(t) = \tilde{E}(t-\tau)e^{i\omega\tau} - \tau\gamma\tilde{E}(t-\tau)e^{i\omega\tau}$$

which lead in the now familiar way to the differential equation $\partial_t E = -\gamma E$ and then to the solution $E(t) = E_0\exp(-\gamma t)$ or $I(t) = I_0\exp(-2\gamma t)$. From this it is seen that the energy lost by the decaying mode in the time τ is $2\gamma\tau^2 I(t)$. This energy may be equated with the expression in Eq. (5.17) and the appropriate decay terms, $\gamma_\pm$, for either mode found:

$$2\gamma_\pm\tau = \oint \frac{I_\pm \delta_\pm^2(\theta,\phi)}{I_+\delta_+^2(\theta,\phi) + I_-\delta_-^2(\theta,\phi)} I(\theta,\phi)\frac{d\Omega}{4\pi} . \tag{5.18}$$

When the expression for $I(\theta,\phi)$ is substituted into this expression the integral becomes intractable without a detailed scattering model to give a form for $\delta_\pm(\theta,\phi)$. Such a model has not been considered in this thesis. The integral is replaced by a simple expression with several parameters using a series of physically reasonable assumptions. The decay terms are

$$\gamma_\pm = \frac{1}{2\tau}\oint \left(\delta_\pm^2(\theta,\phi) + \frac{2\sqrt{I_+I_-}\delta_+(\theta,\phi)\delta_-(\theta,\phi)\delta_\pm^2(\theta,\phi)}{\delta_+^2(\theta,\phi)I_+ + \delta_-^2(\theta,\phi)I_-}\cos(\psi + \epsilon(\theta,\phi))\right)\frac{d\Omega}{4\pi} .$$

The integral is replaced by parameters in the following manner. The first term is simply written as two parameters while the second is manipulated in order to separate uncorrelated pairs (δ_-, δ_+) from correlated pairs (δ_-^2 and δ_+^2) of terms in the integral:

$$\gamma_\pm = \frac{\delta_\pm^2}{2\tau} + \frac{1}{2\tau}\oint 2\left(\frac{I_\mp}{I_\pm}\right)^{\frac{1}{2}} \delta_+\delta_- \left(1 - \frac{\delta_\mp^2 I_\mp}{\delta_\pm^2 I_\pm}\right)^{-1} \cos(\psi + \epsilon_\pm)\frac{d\Omega}{4\pi} \ . \tag{5.19}$$

No specific model for $\delta_\pm(\theta, \phi)$ is used in this thesis but the two functions are taken to be smooth and uncorrelated. For the purposes here the assumption is made that the ratio of the terms $\delta_+^2 I_+$ and $\delta_-^2 I_-$ may be taken as unity for all (θ, ϕ). This assumption reduces the expression for the decay terms to

$$\begin{aligned}\gamma_\pm &= \frac{\delta_\pm^2}{2\tau} + \frac{1}{2\tau}\left(\frac{I_\mp}{I_\pm}\right)^{\frac{1}{2}} \oint \delta_+(\theta,\phi)\delta_-(\theta,\phi)\cos(\psi + \epsilon(\theta,\phi))\frac{d\Omega}{4\pi} \\ &= \frac{\delta_\pm^2}{2\tau} + \frac{1}{2\tau}\mu_\pm\left(\frac{I_\mp}{I_\pm}\right)^{\frac{1}{2}} \cos(\psi + \epsilon_\pm) \ . \end{aligned} \tag{5.20}$$

The regeneration equation now has the scatter loss term $-\tau\gamma_\pm \tilde{E}(t-\tau)e^{i\omega_\pm \tau}$ and the backscatter term, given in Eq. (5.15), included to produce the final regeneration equation

$$\tilde{E}_\pm(t) = \tilde{E}_\pm(t-\tau)e^{i\omega_\pm\tau}\left(1 + \frac{\tau}{2}\left(\frac{\alpha_{0\pm}}{1 + I_\pm(t)/I_{\text{sat}}} - l_\pm\right) - \tau\gamma_\pm + \boldsymbol{\delta}_{\pm\pm}\right) + \boldsymbol{\delta}_{\mp\pm}\tilde{E}_\mp \ . \tag{5.21}$$

where an $\exp(i\omega_\pm\tau)$ term has been inserted to simplify the expression. This term has been seen throughout the manipulations performed so far in this chapter but is not significant as it varies only very slightly from unity and never appears in a difference term where this variation from unity might become important. In order to find the differential equations that govern the field evolution the fast and slow variables are separated according to Eqs. (5.2) and the exponentials collected on the right hand side:

$$\begin{aligned} E_\pm(t) &= E_\pm(t-\tau)e^{i(\mp\phi_\pm(t-\tau)\pm\phi_\pm(t))}\left(1 + \frac{\tau}{2}\left(\frac{\alpha_{0\pm}}{1 + I_\pm(t)/I_{\text{sat}}} - l_\pm\right) - \tau\gamma_\pm + \boldsymbol{\delta}_{\pm\pm}\right) \\ &\quad + \tau\boldsymbol{\delta}_{\mp\pm}E_\mp(t)e^{i(\omega_\mp - \omega_\pm)t \pm i(\phi_+ + \phi_-)} \ . \end{aligned}$$

The definitions made in Eq. (5.3) may be used along with the identification $\phi_\pm(t) - \phi_\pm(t-\tau) = \tau\partial_t\phi_\pm(t)$ and the complex $\boldsymbol{\delta}$ coefficients are expressed in the form $\delta\exp(i\epsilon)$. This reduces the expression to

$$\begin{aligned} \partial_t E_\pm &= \left(\pm i\partial_t\phi_\pm - \gamma_\pm + \frac{1}{2}\left(\frac{\alpha_{0\pm}}{1 + I_\pm(t)/I_{\text{sat}}} - l_\pm\right) + \delta_{\pm\pm}e^{i\epsilon_{\mp\pm}}\right)E_\pm \\ &\quad + \delta_{\mp\pm}E_\mp e^{\pm i\psi + i\epsilon_{\mp\pm}} \ . \end{aligned}$$

The real and imaginary parts lead to the dynamic equations for the field amplitude and the phase difference between the two modes.

$$\begin{aligned}
\mp\partial_t\phi_\pm &= \delta_{\pm\pm}\sin\epsilon_{\pm\pm} + \delta_{\mp\pm}\frac{E_\mp}{E_\pm}\sin(\pm\psi+\epsilon_{\mp\pm}) \\
\partial_t E_\pm &= \left(-\gamma_\pm + \frac{1}{2}\left(\frac{\alpha_{0\pm}}{1+I_\pm(t)/I_{\rm sat}} - l_\pm\right) + \delta_{\pm\pm}\cos\epsilon_{\pm\pm}\right)E_\pm \\
&\quad +\delta_{\mp\pm}E_\mp\cos(\pm\psi+\epsilon_{\mp\pm})\ .
\end{aligned}$$

A set of equations are required which use the irradiances as the dynamic variables. Noting that with $I_\pm = E_\pm^2$ the derivatives of the field amplitudes are written $(\partial_t I_\pm)/I_\pm = 2(\partial_t E_\pm)/E_\pm$ so that

$$\begin{aligned}
\partial_t I_+ &= 2I_+\left(-\gamma_+ + \frac{1}{2}\left(\frac{\alpha_{0+}}{1+I_+/I_{\rm sat}} - l_+\right)\right. \\
&\quad \left. +\delta_{++}\cos\epsilon_{++} + \delta_{-+}\left(\frac{I_-}{I_+}\right)^{\frac{1}{2}}\cos(\psi+\epsilon_{-+})\right) \\
\partial_t I_- &= 2I_-\left(-\gamma_- + \frac{1}{2}\left(\frac{\alpha_{0-}}{1+I_-/I_{\rm sat}} - l_-\right)\right. \\
&\quad \left. +\delta_{--}\cos\epsilon_{--} + \delta_{+-}\left(\frac{I_+}{I_-}\right)^{\frac{1}{2}}\cos(\psi+\epsilon_{+-})\right) \\
\partial_t\psi &= \Omega + \delta_{--}\sin\epsilon_{--} - \delta_{++}\sin\epsilon_{++} \\
&\quad +\delta_{+-}\left(\frac{I_+}{I_-}\right)^{\frac{1}{2}}\sin(\psi+\epsilon_{+-}) - \delta_{-+}\left(\frac{I_-}{I_+}\right)^{\frac{1}{2}}\sin(\psi+\epsilon_{-+}) \qquad (5.22)
\end{aligned}$$

where the decay coefficients, $\gamma_\pm$, are given in Eq. (5.20). These three equations define a 20 parameter model of the oscillations of the internal cavity irradiances of the individual output beams. These internal irradiances are related to the output data stream in two steps. The detection system is calibrated so that the incident power on the PMT is known accurately. This radiant power (W) is related to the internal cavity irradiances (Wcm^{-2}) using the manufacturer's specifications for the transmission of the super mirrors (4 ppm) and using a CCD camera to measure the spot size ($\approx 0.04\,\mathrm{cm}^2$) of the beam on the PMT. Therefore the internal cavity irradiances have to be scaled by $\mathcal{F} = 1.6\times10^{-7}\,\mathrm{cm}^2$ before being compared with the measured output radiant power. Note that all irradiances are measured in Wcm^{-2}.

5.1.7 Investigating Model Parameters

The model that has been developed here has 19 parameters. The data which is analysed later in this chapter has been collected from C-I and has only two harmonics visible above the background noise. Such data constitutes measurement of 14 quantities. The two DC components of the outputs, the peak-to-peak

amplitudes of the fundamental oscillations, the phase shift between these fundamentals and the frequency constitute 6 measurements. Each subsequent harmonic provides four extra measurements, two each from amplitudes and phases. Thus the model has more parameters than quantities that are measured and is over parameterised. Fortunately several parameters may be neglected and others relate to physical quantities that may be measured independently. What follows is a collection of assumptions, used to simplify the model, and measurements of parameters which relate only to the data taken on the 11^{th} and 12^{th} October 1998.

Those parameters that may be removed from the model are discussed first. The forward scatter fields are expected to differ only very slightly in phase from the forward field as the scatterers are assumed to be small relative to the lasing wavelength. This sets $\epsilon_{\pm\pm} = 0$ so that the $\delta_{\pm\pm}$ terms are lost from the phase evolution equation. Next the $\delta_{\pm\pm}$ and $\delta_{\pm}^2$ terms are absorbed by the loss gain difference so that only the saturated gain term governs the ring-up/ring-down behaviour. Further simplification is achieved by shifting the origin of time thereby reducing the degrees of freedom by 1 in the phase subspace. This is equivalent to selecting one phase parameter, say ϵ_{+-}, and measuring the other phases relative to this. Thus the transformation $\psi + \epsilon_{+-} \rightarrow \psi$ is made. This last step does not actually reduce the number of parameters involved in the fitting procedure as an extra parameter is needed to locate the start of the time base when a synthesised waveform is compared with experimental output. However, the removal of ϵ_{+-} from the equations signifies that the parameter is surplus to the physical description of the laser. The equations that model the ring laser output now take the form:

$$\begin{aligned}
\partial_t I_\pm &= 2I_\pm \left(\frac{1}{2\tau}\mu_\pm \left(\frac{I_\mp}{I_\pm}\right)^{\frac{1}{2}} \cos(\psi + \epsilon_\pm - \epsilon_{+-}) + \frac{1}{2}\left(\frac{\alpha_{0\pm}}{1 + I_\pm/I_{\text{sat}}} - l_\pm\right) \right. \\
&\quad \left. + \delta_{\mp\pm}\left(\frac{I_\mp}{I_\pm}\right)^{\frac{1}{2}} \cos(\psi + \epsilon_{\mp\pm} - \epsilon_{+-}) \right) \\
\partial_t \psi &= \Omega + \delta_{+-}\left(\frac{I_+}{I_-}\right)^{\frac{1}{2}} \sin\psi - \delta_{-+}\left(\frac{I_-}{I_+}\right)^{\frac{1}{2}} \sin(\psi + \epsilon_{-+} - \epsilon_{+-}) \,. \qquad (5.23)
\end{aligned}$$

The phase parameters are rewritten in forms which lend themselves to more straightforward interpretation. The phase difference between the two backscatter fields, $\epsilon_{-+} - \epsilon_{+-}$, is replaced by δ. Similarly the phase difference between the two scatter fields is renamed δ_μ with ϵ_μ relating the phases of the backscatter and scatter loss fields. This last step is achieved by using the relations $\epsilon_+ - \epsilon_{+-} = \epsilon_\mu$

and $\epsilon_- - \epsilon_{+-} = \epsilon_\mu + \delta_\mu$. The dynamic equations are finally written in full as

$$\begin{aligned}
\partial_t I_+ &= 2I_+ \left(\frac{1}{2\tau}\mu_+ \left(\frac{I_-}{I_+}\right)^{\frac{1}{2}} \cos(\psi + \epsilon_\mu) + \frac{1}{2}\left(\frac{\alpha_+}{1 + I_+/I_{\text{sat}}} - l_+\right) \right. \\
&\quad \left. + \delta_{\mp\pm} \left(\frac{I_-}{I_+}\right)^{\frac{1}{2}} \cos(\psi + \delta) \right) \\
\partial_t I_- &= 2I_- \left(\frac{1}{2\tau}\mu_- \left(\frac{I_+}{I_-}\right)^{\frac{1}{2}} \cos(\psi + \epsilon_\mu + \delta_\mu) + \frac{1}{2}\left(\frac{\alpha_-}{1 + I_-/I_{\text{sat}}} - l_-\right) \right. \\
&\quad \left. + \delta_{\mp\pm} \left(\frac{I_+}{I_-}\right)^{\frac{1}{2}} \cos\psi \right) \\
\partial_t \psi &= \Omega + \delta_{+-} \left(\frac{I_+}{I_-}\right)^{\frac{1}{2}} \sin\psi - \delta_{-+} \left(\frac{I_-}{I_+}\right)^{\frac{1}{2}} \sin(\psi + \delta) \; . \qquad (5.24)
\end{aligned}$$

These equations have been used to perform the numerical calculations in this chapter. Some of the parameters have been measured directly. The orbit time, τ, is simply the inverse of the free spectral range which is measured to high accuracy; for the data considered this was $\nu_{\text{fsr}} = 86.488\,\text{MHz}$. A thorough analysis of saturation irradiance has been performed by P W Smith [105] for the 632.8 nm He-Ne laser. The relevant results from this work relate operating pressure (Torr) with saturation irradiance (Wcm^{-2}) according to $I_{\text{sat}} = -0.3162 + 6.3738P + 0.7595P^2$. The pressure during the collection of the data examined here was 1.04 Torr giving a saturation irradiance of $7.1\,\text{Wcm}^{-2}$.

All the non-oscillating gain and loss is governed by $\alpha_{0\pm}$ and $l_\pm$. For a low level of waveform distortion the oscillating terms add little overall gain/loss and hence the ring-down and output power may be used to determine $\alpha_{0\pm}$ and $l_\pm$ to good approximation. The cavity ring-down time is measured by switching off the rf supply signal and monitoring the PMT output with a digital storage oscilloscope. The decay time is found to be $\tau_{1/e} = 14.4 \pm 0.6\,\mu\text{s}$ and is the same for both the CW and CCW beams to within the uncertainty. This gives the loss for each cavity as $l_+ = l_- = 6.9 \times 10^4\,\text{s}^{-1}$. By neglecting the oscillating terms the dynamic equations and steady state solution are

$$\partial_t I_\pm = I_\pm \left(\frac{\alpha_{0\pm}}{1 + I_\pm/I_{\text{sat}}} - l_\pm\right) \quad \Rightarrow \quad I_{\text{ss}} = \left(\frac{\alpha_{0\pm}}{l_\pm} - 1\right) I_{\text{sat}} \qquad (5.25)$$

where I_{ss} is the steady state irradiance related to the output power I_0 by $I_{\text{ss}}\mathcal{F} = I_0$. The small signal gain is then approximated by $\alpha_{0\pm} = (1 + I_0/\mathcal{F}I_{\text{sat}})l_\pm$ which provides a useful starting value. The parameters $\mu_\pm, \delta_{\pm\mp}, \delta, \delta_\mu, \epsilon_\mu$ and $\alpha_\pm$ are free and can only be determined by fitting to experimental data. These parameters relate to scatter amplitudes, scatter phases and small signal (unsaturated) gains respectively. It is important to note that $\mu_\pm$ and $\delta_{\pm\mp}$ describe the fluctuations

in the amplitudes of the scatter fields, not the overall amplitudes. The zero frequency components of the scatter fields have been absorbed into the total cavity loss terms $l_{\pm}$ and the cavity gain terms $\alpha_{\pm}$.

5.2 Single Beam Data Collection

The model described thus far specifically describes the separate irradiance fluctuations of each counter-propagating beam. In order to compare theory with experiment data needs to be collected from each beam simultaneously using two photomultiplier tubes. To this end I have constructed a new data gathering system which allows careful calibration and hence provides an accurate measure of the output irradiances. Also, symmetric properties of the system have been utilised to minimise systematic errors. In this manner meaningful comparison between the irradiances is achieved which is subsequently compared with the predictions of the model.

5.2.1 Experimental Arrangement

Previous data collection has simply logged the current oscillations from a PMT while either single or combined beam output is incident. From those records frequency estimates were made. However, if meaningful comparison is to be made between theory and experiment then the output needs to be calibrated and collected in a form appropriate for the model. The experimental arrangement used to achieve this is shown in Fig. 5.5. The model provides a description of the behaviour of the separate irradiances for the counter-propagating beams. Hence two identical arrangements have been used to collect data from each beam. The mirror configuration of the cavity has been changed from $\begin{smallmatrix} F & C \\ C & F \end{smallmatrix}$ to $\begin{smallmatrix} F & F \\ C & C \end{smallmatrix}$ with the two southern mirrors curved. The cavity is thus symmetric about a north-south axis so that beam spot sizes are equal on each photo-cathode. The symmetry of the detection system is further increased by biasing the photomultiplier tubes PMT1 and PMT2 with the same stabilised power supply. The calibration is performed immediately before and after data collection with all voltages kept constant. These steps ensure that the calibration is valid for a particular data set but for that data set only. It is necessary to go to these lengths as the output from a PMT is strongly dependent on the bias voltage; the output signal is scaled by a factor of 2 when the bias voltage is changed from -800 V to -850 V. The 632.8 nm interference filters F1 and F2, the polarisers P1 and P2 and the photomultiplier tubes PMT1 and PMT2 are purchased from the same production runs (same batch number). The distances between optical components are equal and the exit windows are anti-reflection coated and angled to avoid the output beam being reinjected into the cavity and providing coupling between the two modes that is not included in the model.

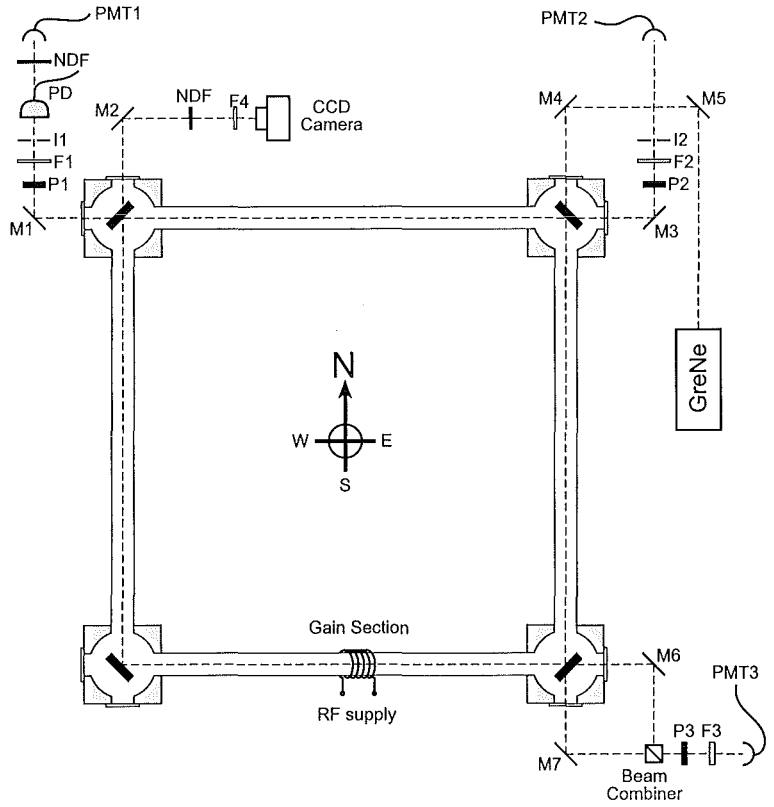


Fig. 5.5: The experimental arrangement used to collect simultaneous single beam profiles is illustrated. Note the symmetry of the data collection portions about a N-S axis.

Background noise which arises partly from the ambient but mostly from the glow of the discharge is reduced in several ways before arriving at the PMT. First it is noted that the output beam is highly polarised with the axis of polarisation perpendicular to the plane of the lasing path. A polariser is placed in the beam path and oriented to allow the maximum lasing power through. This attenuates the discharge background, which has random polarisation, by a factor of approximately 2. Next the narrow spectral properties of the laser beam are exploited using an interference filter with passband centred on to 632.8 nm. Finally the spatially broad background field is attenuated by passing the beam through an iris. The result is that only the spontaneous emission from the lasing transition with the proper polarisation and which is directed along the lasing path will contribute to the background noise in the detector. Note that only a fraction of this noise contributes to the quantum noise considered in Chapter 4. In order to contribute to the phase of the lasing cavity mode the spontaneous emission must be resonant with the mode and this resonance is much narrower (of the order of several kHz) than the atomic transition ($\sim$1.5 GHz).

The cavity alignment is discussed in Chapter 2 and is performed before any other optical elements are added. The alignment system is left in place so that in the

event that one of the free standing mirror holders is moved the cavity may be returned to its original configuration. In practice this safety measure has not been used. An entirely independent optical arrangement is used for the power servo-mechanism. This is shown in the south-east corner. The combined output power of the two modes is used to keep the total output (and hence input) power constant. Further diagnostics are provided using the CCD camera shown. This camera may be used to look at the output beam when lasing at low or high power by using a neutral density filter (NDF) and variable exposure time to avoid saturation of the CCD.

5.2.2 Calibration Procedure

Each of the output channels is calibrated separately. The output current from each PMT is passed through a transimpedance amplifier and through an anti-alias filter before being converted to a digital signal. The entire detection sequence is calibrated in a single procedure thereby giving a constant which converts the integers from the digitisation process into radiant power. Ideally each electronic component in this process would be thoroughly characterised allowing assessment of the performance of each unit. Those culprits which do not perform up to required specifications may then be identified and replaced. This is not feasible with the time and manpower available and hence the more simple approach has to be adopted but which requires a fresh calibration for each data collection session.

The calibration procedure comprises two parts. Initially the calibration device[2] and the detection system have to agree on the zero point. When operating in single longitudinal mode the output power is low and it has to be ascertained whether the background discharge component is significant. To determine this the laser is operated with the discharge on but just below the threshold required to support lasing. Both the photodiode (for the Newport power meter) and the PMT detect no change when the beam alternatingly open and occulted. This type of synchronous detection is a highly sensitive method of identifying small changes in the zero frequency component of a noisy signal. The increase in the discharge brightness required to give strong lasing is only very slight. It is deduced that the contribution from the discharge has been reduced sufficiently to be neglected. With the beams occulted the transimpedance amplifier and the power meter are set to provide a zero measurement.

The power meter requires an output power somewhat greater than that provided by single mode lasing in order to provide a satisfactory reading. Thus the calibration is executed while the lasing is bright and hence requires the beam to be attenuated before reaching the PMT. An OD2 NDF is chosen and is calibrated separately. The arrangement of components is illustrated on the north-west cor-

[2] An 820 series Newport analogue laser power meter was used in conjunction with a model 818-SL silicon photodetector. The power meter has a typical resolution of 1% of full range which for these measurements is 1 nW. The instrument has three calibration settings one of which fortuitously coincides with the He-Ne lasing wavelength at 632.8 nm.

ner in Fig. 5.5. With the NDF in place a data record is taken with several short interruptions where the photodiode (PD) is introduced and a power reading taken. A typical calibration run is provided in Fig. 5.6 where it can be seen that the two different channels have similar calibration constants.

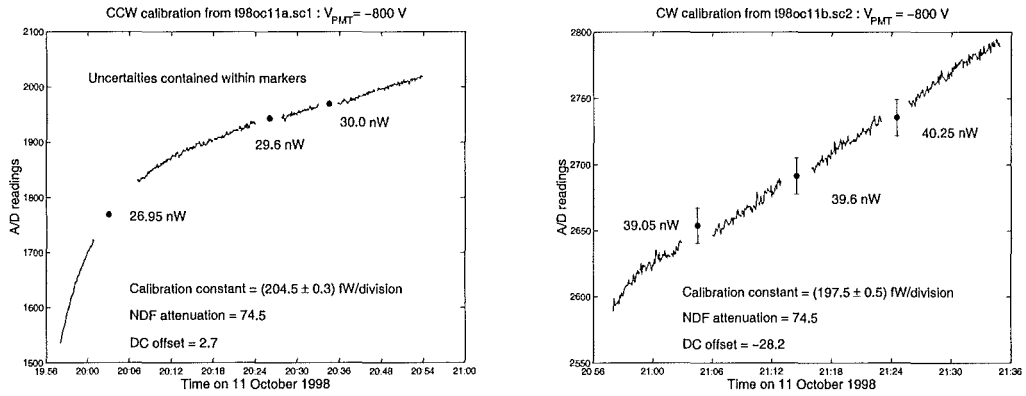


Fig. 5.6: These two plots illustrate the results of the calibrations taken on 11 October 1998. The intervals where the photodiode intercepts the beam to the PMT are clear and in each interval the power meter measurement is shown. The calibration constant is varied until these power measurements are located properly in the PMT record. This method of varying the calibration constant to give best fit also provides the error bars.

As discussed in §2.9 the transverse mode which receives the greatest gain is not necessarily $TEM_{0,0}$ and may vary with different geometries and discharge dynamics. The data considered here was taken under operating conditions where $TEM_{0,2}$ and $TEM_{5,1}$ received roughly equal gain; it is possible to support both modes simultaneously. For the above calibration to be valid for low irradiances (as in single mode operation) it is necessary that the same transverse mode oscillate whatever the output irradiance of the mode is. This requirement follows from the variation of sensitivity of the photo-cathode of the PMT across its surface. Transverse modes have different spatial power distributions and spatial variations in the photo-cathode efficiency will give a different calibration constant for each mode. Often, as the power to the discharge is increased, a second transverse mode oscillates thus invalidating the calibration procedure. The solution has been to adjust the coils of the rf coupling to a position were this effect ceases and only $TEM_{5,1}$ lases whatever the operating power level.

5.3 Comparing the Model with Experimental Results

In this chapter a model has been formulated to describe single beam irradiance oscillations. A detection scheme has been described to collect data in a form which is described by the model. Initially a comparison between the predicted and real waveforms was intended. The free parameters would be varied to achieve

best fit. To this end a program has been written in MATHEMATICA® to solve the differential equations numerically and to optimise the fit by varying the parameters. Unfortunately, despite MATHEMATICA® containing an excellent numerical package, the numerical analysis is presently too computationally expensive to be performed in the time available; one second of data has required several days to analyse. The performance of the model has to be assessed in other ways.

In each of the sections that follow realizable methods of investigating the model's capacity to describe ring laser performance are described. The first approach requires data collected with the power to the discharge set at several distinct levels. The model is fitted to the waveform found at each power level by varying the model parameters until optimal fit is found in a least squares sense. The evolution of these parameters with time-averaged combined beam output power is found thus providing an insight into the stability of the fitting method and a check that the physically meaningful parameters are behaving as expected. A further investigation relies upon an important observation that I made early in my experimental work. I noticed that when a small oscillating voltage is applied to the piezo-driver, located behind the south-east mirror, and the frequency of oscillation varied it is possible to make the ring laser output frequency and the piezo frequency lock together. In order to model this phenomenon the parameters in the model that relate to the phase and amplitude of scatter fields are modulated and the resulting beat frequency found numerically. It is not surprising that both the ring laser and the model lock to the piezo modulation frequency as they both represent nonlinearly coupled oscillators. However, a comparison between the observed locking profile and that predicted by the model provides further verification of the model's validity. The final inspection of the model's behaviour is provided by varying the rotation rate of the ring laser numerically. Although such experiments are not implemented locally other groups provide appropriate results and observations which compare well with the predictions of the model.

5.3.1 Parameter Versus Power Dependence

Under ideal conditions, where time and computing power are not limiting, the performance of the model could be assessed in a straightforward manner. Dual data streams could be recorded for a long period of time and the model fitted to small segments of the data, for example every second. Occasional duty cycles could be automatically exercised to measure those parameters that lend themselves to direct measurement and a detailed account of how parameters affect output frequency could be established. This is not currently possible. The numerical analysis of the ten data segments considered here required a week of processing time.

An active approach was taken to identify whether the model behaves in a reasonable manner. I have selected a single operating parameter which may be varied to give several distinct output waveforms. This parameter is the set point voltage on the power servo-mechanism. While unlocked the power was set to a very low

level and data collection started. Every 30 s the power level was increased and single mode operation verified using an rf spectrum analyser which monitored the servo PMT output. This process was repeated until a free spectral range signal was observed signifying multimode operation at which point data collection was stopped. The data set analysed here was recorded on October 12$^{\text{th}}$ 1998 and is stored in the two files `t98oc12a.sc1` and `t98oc12a.sc2`, one for each beam. Ten separate power levels were found which provided single mode lasing. Using a simplex method waveforms have been fitted to the output data (this process requires little processing time). The waveforms are each formed from a fundamental and two harmonics, specifically:

$$\begin{aligned} I_{-}^{*}(t) &= A_0 + \sum_{j=1}^{3} A_j \sin(2\pi j f_b t + \phi_j + \phi) \\ I_{+}^{*}(t) &= B_0 + \sum_{j=1}^{3} B_j \sin(2\pi j f_b t + \theta_j + \phi) \end{aligned} \tag{5.26}$$

where the $*$ signifies an experimentally determined waveform. The ten data segments are described by the parameters provided in Table 5.1. This table includes the rms fitting error per data point, Δ^{*}_{rms}, which shows that the parameters given in the table provide a good description of the actual output waveforms. The

	1	2	3	4	5	6	7	8	9	10
A_0	40.48	64.33	93.15	136.05	163.96	204.38	269.66	344.38	436.41	544.95
A_1	10.58	17.59	25.95	37.92	45.80	56.82	73.64	90.88	109.29	127.38
A_2	0.06	0.27	0.73	1.45	1.97	2.96	4.62	6.91	9.64	12.65
A_3	0.05	0.05	0.06	0.03	0.07	0.03	0.22	0.26	0.57	1.14
B_0	22.92	44.47	70.72	110.01	135.60	172.35	231.25	298.43	380.84	477.95
B_1	10.69	18.30	27.44	40.61	48.82	59.98	76.32	92.46	109.15	124.89
B_2	0.20	0.33	0.72	1.37	1.83	2.49	3.35	4.34	5.48	7.23
B_3	0.04	0.03	0.06	0.05	0.06	0.03	0.15	0.17	0.06	0.21
$\theta_2 - 2\theta_1$	5.135	5.547	5.485	5.523	5.530	5.512	5.429	5.278	5.050	4.832
$\theta_3 - 3\theta_1$	2.832	2.412	3.179	3.139	3.464	3.211	2.768	2.612	3.055	3.179
$\phi_1 - \theta_1$	3.415	3.396	3.375	3.336	3.309	3.282	3.251	3.229	3.212	3.201
$\phi_2 - 2\theta_1$	3.262	1.563	1.676	1.615	1.560	1.526	1.489	1.464	1.436	1.412
$\phi_3 - 3\theta_1$	4.354	3.364	2.902	3.014	3.429	3.211	4.167	4.087	4.181	4.650
f_b	72.824	72.774	72.684	72.512	72.380	72.147	71.710	71.167	70.495	69.778
Δ^{*}_{rms}	1.195	1.282	1.453	2.140	2.311	2.217	2.662	2.991	3.065	3.986

Table 5.1: These parameters represent the measured waveforms of the counter-propagating beams according to Eqs. (5.26). The rms fitting error per data point, Δ^{*}_{rms}, is included to show the quality of the fitting procedure. All phases are estimated at a point in time where θ_1 takes a specific value. The point is chosen to be where $\theta_1 = 0$ without any loss of generality.

waveforms $I_{\pm}^{*}(t)$ are used by MATHEMATICA to find those model parameters that provide best fit. A given set of model parameters is tested for quality of fit by solving the dynamic equations (Eqs. (5.23)) with those values and finding the deviation of $I_{\pm}(t)$ from $I_{\pm}^{*}(t)$. This deviation is most quickly determined by choosing a sequence of evenly spaced sample points and determining both $I_{\pm}(t)$ and $I_{\pm}^{*}(t)$

at those points. The rms deviation per sample point, $\Delta_{\rm rms}$, is then a measure of quality of fit. The problem may be simply stated as the search for the global minimum of $\Delta_{\rm rms}(\alpha_+, \alpha_-, \mu_+, \mu_-, \delta_{+-}, \delta_{-+}, \delta, \epsilon_\mu, \delta_\mu)$. The fitting routine attempts to minimise $\Delta_{\rm rms}$ by varying the parameters using a *path of steepest descent* algorithm in parameter space which normally converges towards the optimum solution faster than simplex algorithms. Notice that the actual experimental data is not used to perform the model parameter estimation. This enables a considerable decrease in the convergence time of the algorithm. The raw data has several spectral features with a broadband base noise. If this noise is passed through to the fitting procedure then the minima in parameter space are not smooth functions of the parameters but will have a noisy profile. Using the fitted waveforms $I^*_\pm(t)$ removes the noise from the system allowing faster convergence.

The numerical values of the model parameters at optimal fit are provided in Table 5.2. Also reported in this table are the values for three separate fitting parameters $I_+(0)$, $I_-(0)$ and ϕ. The first two of these represent the initial values for the field irradiances which are required for the numerical solution of the dynamic equations and the third is the shift between the time frames of $I_\pm(t)$ and $I^*_\pm(t)$ and is presented as a phase angle (see Eqs. (5.26)). These three parameters are included in the table for completeness. Finally the quality of the fit between $I_\pm(t)$ and $I^*_\pm(t)$ is indicated by Δ which is expressed as a percentage. These values are calculated from the rms fitting error per data point which is scaled by $A_0 + B_0$ in order to give a fractional error per point. From the values for Δ it may be seen that the first and last points have achieved slightly lower convergence than for the other data segments; otherwise the quality of fit is roughly equal for all data segments. The 9 physically meaningful parameters are displayed in Fig. 5.7 where

	1	2	3	4	5	6	7	8	9	10
ϵ_μ	1.978	1.993	2.041	1.860	1.870	1.841	1.890	1.913	1.901	1.917
δ_μ	4.262	4.255	4.179	4.254	4.254	4.258	4.207	4.199	4.206	4.193
δ	0.086	0.069	-0.081	-0.188	-0.178	-0.209	-0.277	-0.201	-0.074	-0.093
μ_+	73.40	64.36	49.31	67.48	64.57	64.67	58.07	60.50	69.87	65.46
μ_-	313.52	313.98	307.35	363.74	366.52	365.89	387.00	405.32	395.14	393.26
δ_{+-}	76.99	81.78	81.04	114.73	117.61	119.76	133.78	145.96	145.32	149.65
δ_{-+}	61.28	61.30	61.42	54.30	53.51	52.00	48.56	43.76	36.92	31.80
$\alpha_+ - l$	1.43	4.64	8.48	15.67	17.17	19.91	25.17	29.16	33.41	39.56
$\alpha_- - l$	0.0023	1.57	2.23	2.31	4.18	6.19	8.73	13.27	19.40	25.58
$I_+(0)$	18.63	38.36	61.01	89.12	112.39	141.92	193.88	258.03	332.82	422.99
$I_-(0)$	42.11	65.99	96.64	149.49	178.78	226.11	298.56	376.92	477.11	592.99
ϕ	5.86	5.95	5.92	5.75	5.81	5.79	5.81	5.89	5.89	5.90
$\Delta(\%)$	0.329	0.275	0.275	0.270	0.263	0.265	0.257	0.269	0.281	0.314

Table 5.2: The results of optimising model parameters are reported here. The nine physically meaningful parameters are reported first with three fitting routine specific parameters reported next. Note that $l = 6.9 \times 10^5$. The quality of fit is given last and is presented as a percentage.

they are separated into scatter phases, scatter amplitudes and small signal gains. These results are presented against the sum of the time averaged power for both beams, $A_0 + B_0$. The measured waveforms considered here were all collected within a 5 minute period. The output frequency was noted to be particularly

stable before and after the data was gathered and hence parameter drift during data collection is assumed to be minimal. Segment 1 shows little harmonic content (parameters A_2, A_3, B_2, B_3) but this becomes pronounced as the power level increases. A key feature is the spread of output frequencies which span 3 Hz. A strong test of the model then is to see whether the parameter values at optimal fit are appropriately stable while still being able to describe this variety of output waveforms.

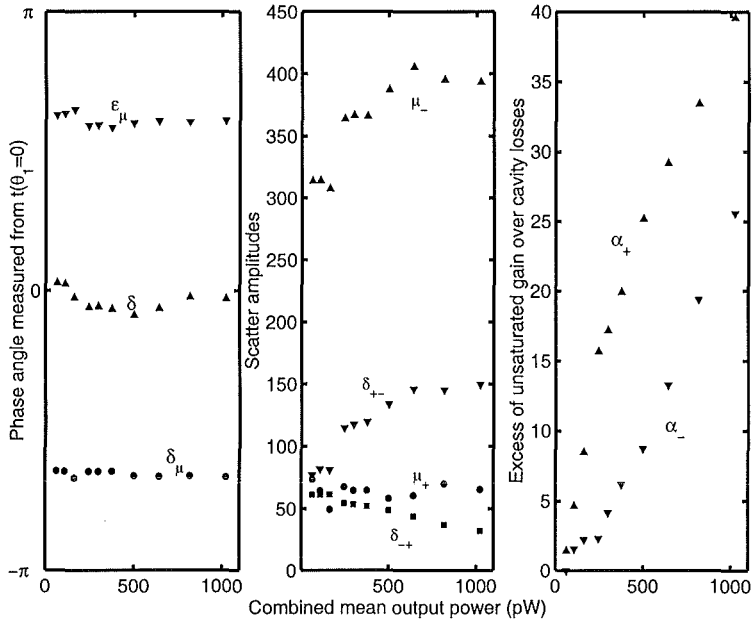


Fig. 5.7: The values of the parameters which provide optimum fit for the ten data segments are presented here. The three plots refer to scatter phases, scatter amplitudes and gain factors from left to right. The parameters are presented against the sum of the time averaged power for both beams, $A_0 + B_0$.

Fig. 5.7 shows that the parameters that relate to scattering remain reasonably constant for the different power levels. Also the small signal gain parameters $\alpha_\pm$ are roughly linearly related to $A_0 + B_0$. This dependence has previously been predicted in Eq. (5.25). Emphasis must be placed on the fact that the theoretically expected Sagnac frequency, 68.95 Hz, has been assumed throughout yet the model has had little difficulty in providing waveforms with frequencies spread over a few Hz and which are higher than the expected frequency. Further, the waveforms have varying levels of distortion. Thus the model has precisely fitted a range of different waveforms yet the parameter values required to achieve these fits do not change appreciably from one waveform to the next. Although the changes are significant enough to cause a considerable frequency shift the smooth dependence on combined beam power indicates that the estimates are reliable.

Although the results presented here constitute a considerable success for the model an explanation is required for why the parameters do not remain com-

pletely steady and why these similar sets of parameter values can provide such different output frequencies. One explanation for the results is provided. It is assumed that the spread of frequencies are due to the spread of scatter amplitudes. A mechanism which affects a change in scatter amplitudes when power levels are altered is required. Such a mechanism is provided by dispersion. The changing power levels cause the refractive index in the discharge to change. In his thesis [128] Ziyuan Li describes dispersion in the gain medium and how this directly affects the beat frequency of a ring laser by altering the optical path length of the cavity. It is concluded that for the typical (very low) values of gain required to make the high Q cavities lase in single mode the effect of dispersion on the beat frequency is negligible compared with the typical frequency excursions observed (see Fig. 1.1). However, the model described here includes other features which are affected by small path-length changes.

The parameters which relate to scatter fields are therefore affected by dispersion. Recall that the parameters μ_+, μ_-, δ_{+-} and δ_{-+} relate to the amplitudes of the fluctuations in the scatter fields and not the total amplitudes of these fields; the zero frequency component of the scatter field being absorbed by the losses $l_\pm$. The overall phase of the scatter fields will only be slightly affected by dispersion. However, each field is the sum of a finite number of components from individual scatter centres. These components have random phases but are not numerous enough to provide a uniform loss so that the above parameters are small but nonzero. Dispersion affects the phases of all the scatterers very slightly so that the total scatter losses remain almost unchanged while the fluctuations in these losses may be significant. This describes the behaviour reported in Fig. 5.7. The large shift in frequency is simply a result of the critical dependence on these parameters.

5.3.2 Cavity Perimeter Modulation

The observation that the ring laser output frequency may be locked to an externally applied perimeter modulation provides a further test of the model's capacity to describe ring laser behaviour. Fig. 5.8 illustrates the locking process. While unlocked and oscillating at 72.60 Hz an sinusoidal voltage of 2 $V_{p\text{-}p}$, zero offset and frequency f_m is applied to the piezo. The manufacturer's calibration of the piezo stack gives a constant of 3.2 nm/V; the inherent piezo hysteresis is not significant at the small voltages used here. Thus the amplitude of the mirror oscillations is of the order of $\lambda/200$. With the modulation frequency initially far below 72.60 Hz the ring laser output frequency, f_b, is monitored with an audio spectrum analyser. Also a Lissajous pattern is formed on an oscilloscope with the ring output (f_b) and piezo driving voltage (f_m) on orthogonal axes. While increasing f_m in discrete steps f_b is recorded. The frequency resolution attainable when determining f_b is limited by drift in the ring laser; this accounts for the scatter of the results in Fig. 5.8. The Lissajous pattern provides a good method for determining f_b when near locking. The rotation rate of the pattern gives the difference in frequency

between f_b and f_m, the latter is known with high accuracy.

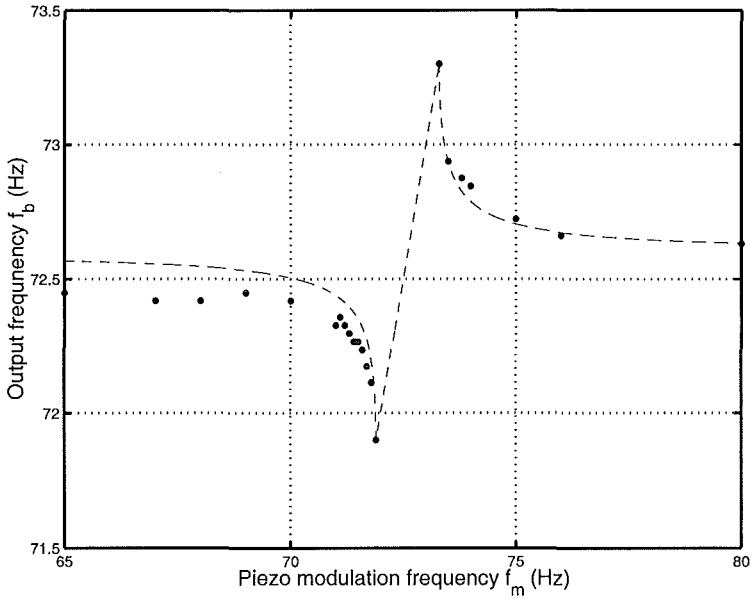


Fig. 5.8: The measured beat frequencies are shown as black circles. The dashed line represents the Adler lock-in profile when the decoupled frequency is 72.60 Hz and the lock band is 1.4 Hz wide.

Key features from Fig. 5.8 ought to be noted. The lock-in band is clear and is symmetric about the decoupled oscillation frequency of 72.60 Hz. The asymptotic frequencies above and below the locking region are not equal. Visualisation of this is aided by the inclusion of the Adler profile in the figure. This profile is evaluated using the concise form

$$f_b = f_m \pm \mathrm{Re}\sqrt{(f_m - f_s)^2 - f_b^2} \tag{5.27}$$

where the + and - refer to the higher and lower frequency regions respectively and f_s is the decoupled frequency. The Adler profile agrees well with the measured frequencies above the lock band but below the lock band the agreement is poor. The asymmetry cannot be attributed to frequency drift because the measurements were made sufficiently quickly to rule this out. Further, the output frequency returned to the initial decoupled frequency when the amplitude of the modulation was reduced to zero.

The applicability of the earlier model of the ring laser is now investigated. Only a limited analysis can be performed here due to the time restrictions imposed by the computation involved in numerical solution of the coupled differential equations. The model is modified to account for the mirror oscillation in the following manner. Each of those parameters that relate to scatter amplitudes and phases are modulated symmetrically about a specified central value. These central values are chosen from the fitted solutions found in Table 5.2. The parameters for

segment number 7 are used as they represent the best fit for all the segments of the data set t98oc12a. The modulation is completely synchronous for all parameters reflecting the fact that as the piezo passes through its central position all the parameters are equal to those chosen from Table 5.2. The magnitude of the modulation is estimated separately for the scatter phases and for the scatter amplitudes. The phase modulation is achieved by inserting $\Delta\phi \sin(\omega_m t)$ whenever a δ phase parameter is encountered; the amplitude of the modulation, $\Delta\phi$, is taken as $2\pi/200$. The modulation of the scatter amplitudes is achieved by inserting the term $(1 + A\sin(\omega_m t))$ in front of each scatter amplitude where $\omega_m = 2\pi f_m$. The scatter amplitudes represent the peak to peak fluctuations in the scatter fields and not the overall size of the scattered fields. The average components of the scatter fields were absorbed by the overall loss term $l_\pm$ earlier in the development of the model. It is expected then that the amplitude modulation depth will be a considerable fraction of the total amplitude. The value $A = 0.25$ has been chosen in the calculations that follow. The modified model is now represented by the three coupled differential equations:

$$\begin{aligned}
\partial_t I_+ &= 2I_+ \left(-\frac{\mu_+}{2\tau}(1 + A\sin(\omega_m t)) \left(\frac{I_-}{I_+}\right)^{\frac{1}{2}} \cos(\psi + \epsilon_\mu) + \frac{1}{2}\left(\frac{\alpha_{0+}}{1 + I_+/I_{\text{sat}}} - l_+\right) \right. \\
&\quad \left. + (1 + A\sin(\omega_m t))\delta_{-+} \left(\frac{I_-}{I_+}\right)^{\frac{1}{2}} \cos(\psi + \delta + \Delta\phi \sin(\omega_m t)) \right) \\
\partial_t I_- &= 2I_- \left(-\frac{\mu_-}{2\tau}(1 + A\sin(\omega_m t)) \left(\frac{I_+}{I_-}\right)^{\frac{1}{2}} \cos(\psi + \epsilon_\mu + \delta_\mu + \Delta\phi \sin(\omega_m t)) \right. \\
&\quad \left. + \frac{1}{2}\left(\frac{\alpha_{0-}}{1 + I_-/I_{\text{sat}}} - l_-\right) + (1 + A\sin(\omega_m t))\delta_{+-} \left(\frac{I_+}{I_-}\right)^{\frac{1}{2}} \cos(\psi) \right) \\
\partial_t \psi &= \Omega + (1 + A\sin(\omega_m t))\delta_{+-} \left(\frac{I_+}{I_-}\right)^{\frac{1}{2}} \sin(\psi) - \\
&\quad (1 + A\sin(\omega_m t))\delta_{-+} \left(\frac{I_-}{I_+}\right)^{\frac{1}{2}} \sin(\psi + \delta + \Delta\phi \sin(\omega_m t)) \; . \qquad (5.28)
\end{aligned}$$

These equations are solved numerically for several values of f_m which are chosen to highlight the lock-in band. Accurate estimates of frequency are required in order to identify fine features of the lock-in profile. This is achieved by noting that the amplitude envelope is modulated at a rate equal to the difference between f_b and f_m. The period of the oscillation in the amplitude envelope is found accurately from a graph of the numerical solution and from this an estimate of f_b is made. Fig. 5.9 illustrates the results of the numerical simulation. The beat frequency when $A = 0$ is known to be $f_s = 71.71$ Hz and this is indicated in the figure by a continuous line. The dashed line indicates the predictions of the Adler theory for the same decoupled frequency and lock-in band. The upper plot clearly shows the asymmetry in the lock-in profile. With f_m initially below the lock band f_b is

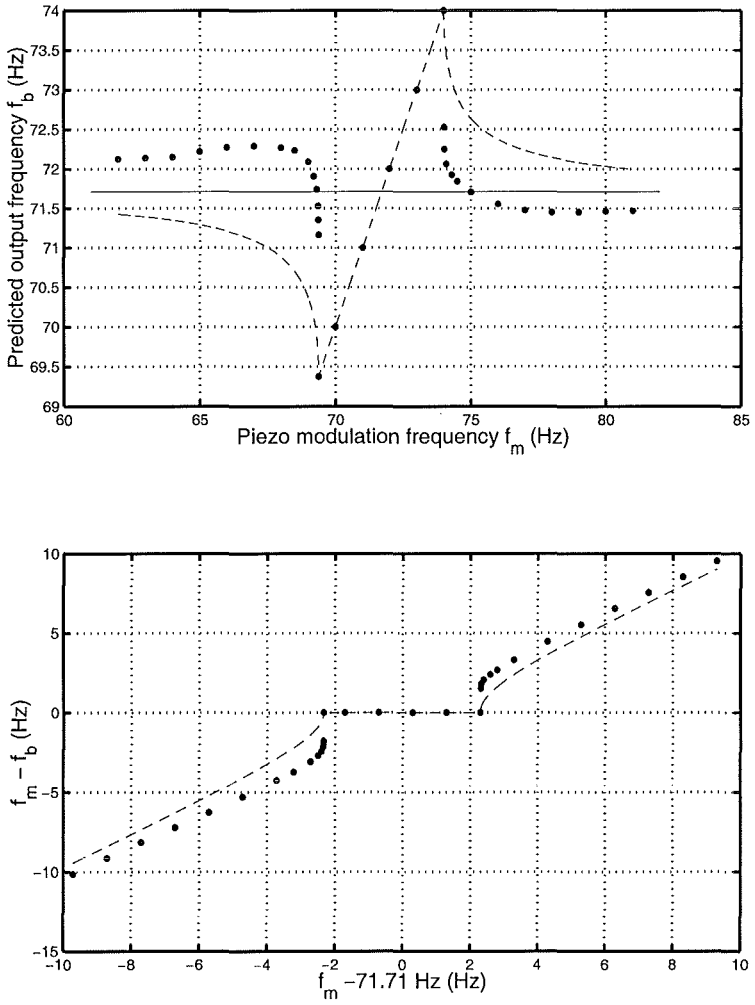


Fig. 5.9: The lock-in profile predicted by the model which has perimeter modulation is illustrated by the black circles. The accuracy to which the predicted frequencies may be stated is smaller than the circles. The solid line indicates the beat frequency in the absence of modulation. The dashed line shows the predictions of Adler theory for the same location and width of the lock-in band.

higher than the decoupled frequency. As f_m approaches locking f_b initially rises but then drops until $f_m = f_b$. This behaviour is not apparent above the lock-in band.

A comparison between the predictions of the model and the measurements indicates a few broad similarities. Both the measured and predicted profiles differ substantially from the predictions of the Adler theory and give asymmetric lock-in profiles. Also the high and low asymptotic frequencies are not equal. Hence the model predicts a type of behaviour which is not described by any other known theory. The simulation performed here uses parameters which were obtained under different conditions from those that were in place when the locking phenomenon was observed. It is only reasonable then to make qualitative conclusions from

the results reported here. A detailed analysis of the locking phenomenon would require continuous sampling of the waveform while the modulation frequency is scanned past the lock-in band. In this way the operating parameters could be found as in §5.3.1 and a direct comparison between the model and the waveform could be made. This investigation may be performed at a later date.

5.3.3 Variation of the Earth's Rotation Rate

This section heading refers to the ease with which ring performance under different rotation rates, Ω, may be investigated using the model. During the time of the work documented here there has been no capacity to vary Ω of any of the Canterbury ring lasers; all lasers are rigidly contacted to the Earth. However, the dependence of the model on Ω may be compared with those of other authors [45] and also the observations of several experimental workers [6,7,9,10,60]. These authors have worked on gyros which are mounted on turntables to allow the study of ring behaviour as Ω is varied. Although the investigation reported here is not exhaustive a considerable qualitative understanding may be obtained by comparing results.

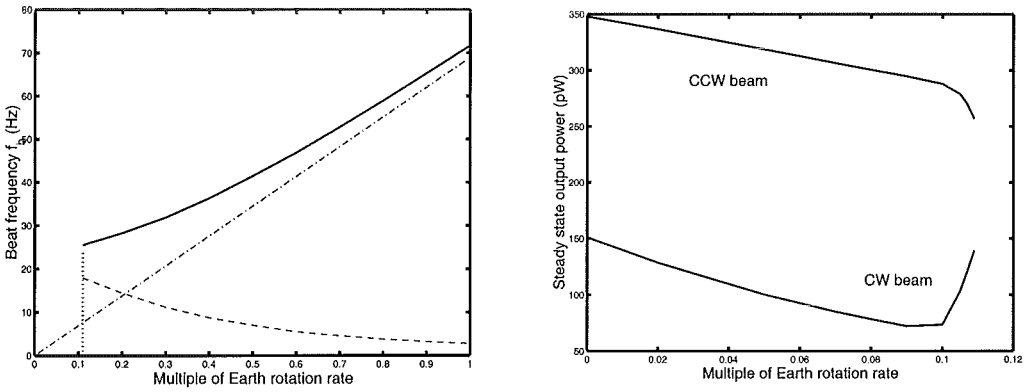


Fig. 5.10: The figure on the left indicates the predicted beat frequency dependence on Ω. The straight dashed line indicates the ideal behaviour while the other dashed line indicates the departure from ideal of the predicted results.

The dynamic field equations (Eqs. (5.23)) are solved numerically using the parameters from segment 7 of t98oc12a.sc* just as in the previous section. This is repeated for a variety of values for the Earth rotation rate, the results are presented in Fig. 5.10. The plot on the left shows the beat frequency which is always above the ideal. This behaviour reflects precisely the prediction made by Etrich et. al. [45, Fig. 8] for conservative coupling. On the right the dependence of the stable beam irradiances are shown in the band from no rotation up to unlocking. Note that although the beams have not unlocked varying Ω has a marked effect on the beam irradiances which reflects the observations made by Aronowitz and coworkers [9] which they refer to as winking.

For the particular parameter values used here the counter-propagating modes unlock at $\Omega = \Omega_L = 0.11\Omega_E$ where Ω_E is the Earth rotation rate. At the point where the ring laser unlocks the solution branches from stable irradiance solutions to oscillating solutions. The oscillating solutions have maximum, mean and minimum irradiances for each beam as a function of Ω as shown in Fig. 5.11. A similar bifurcation arises from the model discussed by Etrich et. al. [45] which is also discontinuous and predicts the same frequency dependence as Fig. 5.10. In a real system such discontinuous dependences are rarely physically reasonable. In the situation considered here it is most likely that quantum noise due to spontaneous emission would round off the corners. This is analogous to the demonstration that quantum noise rounds off the corners of the Adler lock-in band which is due to Cresser et. al. [35].

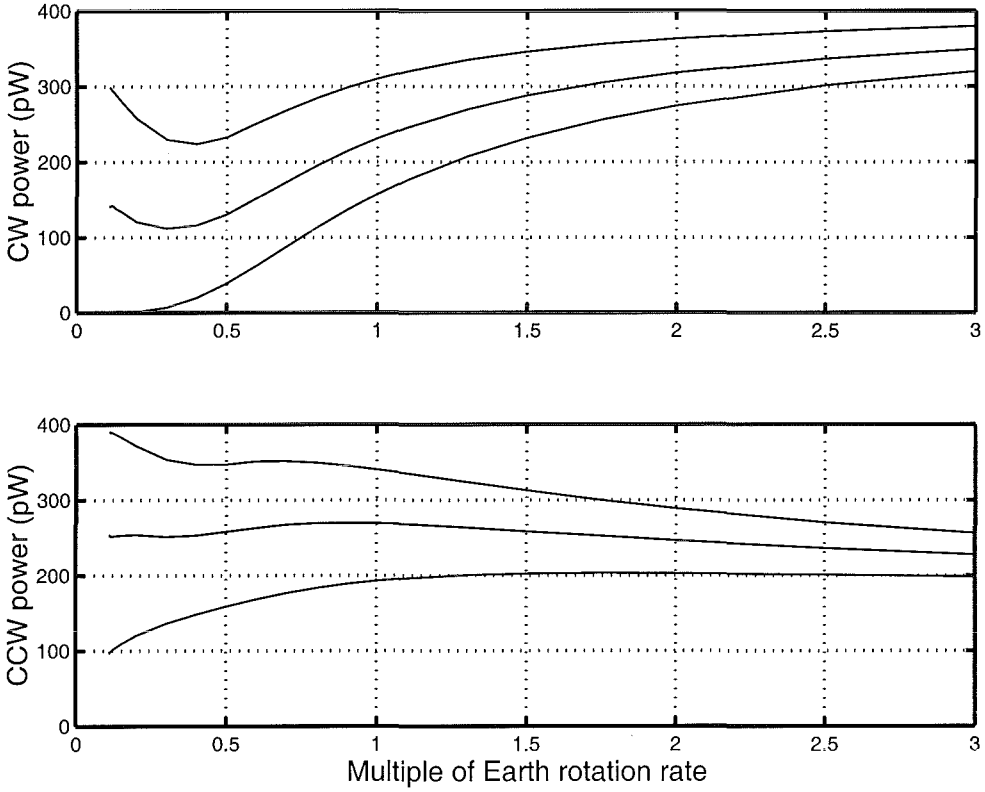


Fig. 5.11: The maximum, mean and minimum irradiances for the CW and CCW modes are illustrated. The oscillations begin at Ω_L. The oscillations in the single beams are still apparent at rotation rates far in excess of Ω_E. As Ω increases the peak-to-peak amplitudes of the oscillations tend to zero.

Huthchings et. al. [60] report a hysteresis effect when they rotated their gyro from below lock-in to above and then back again. Their results are shown in Fig. 5.12 for two different discharge currents. It is possible to reproduce these results using the model currently under consideration. First the following observation is made

which regards the solution of the dynamic equations. Setting Ω slightly below Ω_L and the initial irradiances at some level other than their equilibrium value oscillating numerical solutions are found with a frequency close to that at the point of unlocking. These oscillations decay in amplitude until the equilibrium value is obtained. As Ω is moved closer to Ω_L the persistence of these oscillations becomes longer (many seconds decay time) until finally with $\Omega = \Omega_L$ the oscillations become stable.

The above figures only represent the steady state values for irradiance but the experiment performed by Hutchings et. al. was necessarily performed within a particular finite time frame during which the steady state was not necessarily reached. In Fig. 5.12 the path around the hysteresis effect is indicated. As Ω is increased past Ω_L the gyro unlocks and the beat frequency rises as Ω increases. When Ω is reduced back to Ω_L the oscillations persist with locking only occuring at some rotation rate lower than Ω_L. The model predicts precisely this behaviour. When Ω is reduced below Ω_L the oscillations will require some time to decay to the steady state value. While these oscillations decay Ω continues to reduce and hence oscillations are still observed when $\Omega < \Omega_L$. The model therefore makes a clear prediction that as the scan rate is reduced the point at which oscillations cease will tend towards Ω_L and the hysteresis will vanish as the scan rate tends to zero.

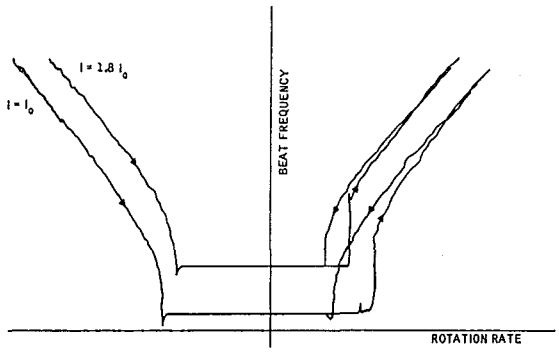


Fig. 5.12: This image comes from the review article by Aronowitz [9] and is originally due to Hutchings et. al. [60]. The apparent hysteresis is clear for the two discharge currents used.

Ziyuan Li [128–130] describes a method for measuring the quality factor of a supercavity. This involves tuning a Fabry-Perot cavity past an injected laser frequency and observing the asymmetric response. As the resonance of the FP cavity moves close to the lasing frequency the field inside the FP cavity builds up. Once past resonance the field decays with a characteristic time defined by the cavity quality factor. While the field decays the resonance of the cavity progresses and a beat between the cavity resonance and the decaying field is observed giving the asymmetric response. Knowledge of the lasing frequency, cavity resonance scan rate and ring down profile allow the cavity Q to be determined. An entirely equivalent experiment could be performed with the lock-in phenomenon. First a precise value of Ω_L could be obtained when increasing Ω above the lock-in threshold. Subsequently, while reducing Ω below Ω_L, a record of the decaying oscillations would be made. This information along with the scan rate of Ω would

provide a second data stream which would complement the steady state unlocked data stream which is routinely obtained. Alternating between the two modes of operation, steady state and transient, would provide extra information. The parameters relating to the amplitude and phase of the scatter fields would be the same for each mode of operation and thus the model would be required to describe two quite different types of behaviour with the same set of parameters. The information arising from such an experiment would place tighter requirements on the performance of the model. If the model satisfied these requirements then this process would undoubtedly determine operating parameters with greater accuracy.

6. A Very Large Ring Laser Gyro

The difference between the experimentally observed and theoretically predicted output Sagnac frequencies is dependent on the size of the ring laser. Scaling rules for ring lasers are discussed by Stedman [108] where it is reported that the improvement in performance may scale as P^2 or P^3 depending on circumstances. The interest is therefore in the performance of large ring lasers ($P{\sim}16\,\mathrm{m}$) relative to the smaller lasers ($P{\sim}4\,\mathrm{m}$) previously in operation. Further information about the effects of frequency drift are provided by the performance of these larger ring lasers.

6.1 Introduction

A very large ring laser (known as G0) has been built and operated in the Cashmere caverns [99] as a prototype for a similar instrument which is to be built for the Forschungsgruppe Satellitengeodäsie. The laser is square, 3.5 m along each edge, and is located on the concrete wall of the cavern. It is believed to be the largest ring laser gyro in the world[1]. The interest in such a large ring laser is due to the scaling up of the output frequency, linearly with linear dimension, and the scaling down of the backscatter mode coupling, quadratically with linear dimension. However, the increase in perimeter introduces the problem that the free spectral range of the instrument scales down with linear dimension. With the longitudinal modes more closely spaced the power servo mechanism is required to be somewhat more precise in order to guarantee single mode operation. Further, with two 6 m curvature mirrors the cavity is only just within the stability region for a $\begin{smallmatrix} \mathrm{F} & \mathrm{C} \\ \mathrm{C} & \mathrm{F} \end{smallmatrix}$ configuration suggesting that alignment is critical unlike the somewhat more forgiving alignment requirements of C-I. With the geometry close to the edge of stability the spot sizes become quite large at various parts of the cavity which magnifies the effects of vignetting of beams by the physical manifold of the cavity. The transverse modes that oscillate provide further information about large ring laser gyros.

This chapter provides a digest of the operating conditions of the laser. The technical details of cavity stability, design, alignment method, operation and

[1] Dunn [42] has anticipated some of the results reported here. This work was only recently reported (September 1998) and refers to the design of a $\sim$ 40 m perimeter triangular ring laser. This ring laser has a low quality factor which translates into a high lock-in threshold of $\sim$ 200 Hz and unlocks easily under the Earth rotation induced Sagnac frequency of $\sim$ 500 Hz.

expected output are described in § 6.2 to § 6.6. A provisional data set is analysed in § 6.7. This data stream is only of the order of a few tens of minutes long as the power servo mechanism was unable to maintain single longitudinal mode operation for longer than this time. Verification that the cavity is oscillating in a single longitudinal mode is achieved by viewing the rf spectrum analyser throughout the data acquisition period to make certain that no fsr beat frequency appears.

6.2 Cavity Stability

As part of the design process for a new laser it is necessary to perform stability calculations in order to determine whether the cavity resonances are stable. Here a novel approach for the determination of a cavity's stability is used. This method is in principle the same as any other but leads to an elegant graphical presentation of cavity stability.

An insight into the method can be obtained by considering a Gaussian beam propagating within a rectangular cavity defined by four plane mirrors. Starting with the waist on a mirror the beam propagates from one mirror to the next with the spot size increasing as it progresses. By the time it returns to the starting point there is a nett spot size increase. This process continues for each orbit and the beam diverges; this is true for all initial spot sizes and hence the plane mirror cavity cannot support a Gaussian mode. The insertion of a focusing element is required to compensate for this divergence and thereby allow the beam to return to its initial spot size and curvature after a single orbit.

It is easily shown that an arbitrary ABCD matrix which transforms a beam at its waist back to its initial spot size will also return it to zero curvature if the matrix satisfies the requirement that A=D. All of the possible configurations of the square cavity (using two flat mirrors and two of equal curvature) have an axis of symmetry and any resonant mode will have a waist on this axis. The test beam is chosen so that it arises from the point where this waist will be found if a resonant mode does exist. The symmetry of the cavity will provide a symmetric sequence of translations and lensing matrices and it is a simple matter to show that such a symmetric sequence will satisfy the A=D criterion. Thus any beam that starts from its waist at the point where a resonant mode will have its waist and which returns to its original spot size after a single orbit will automatically have zero curvature and will therefore trace the path of the resonant mode exactly.

It can be shown that an arbitrary ABCD system that satisfies the A=D criterion (an ABCA system) and which receives a beam at its waist as input will give a monotonically increasing output spot size as the input spot size decreases. Now the spot size of an infinitesimally small spot diverges infinitely rapidly as it propagates and hence no system with a translation component will be able to focus the spot back to its original size. If an extremely large spot, which is

effectively a plane wave, is not shrunk by the system then there is no intermediate spot size that has the property of returning to its initial spot size and the cavity is unstable. However if a plane wave is focused then such an intermediate spot size must exist and the cavity has a stable resonance. Further, the monotonic property noted above shows that this intermediate spot size is unique.

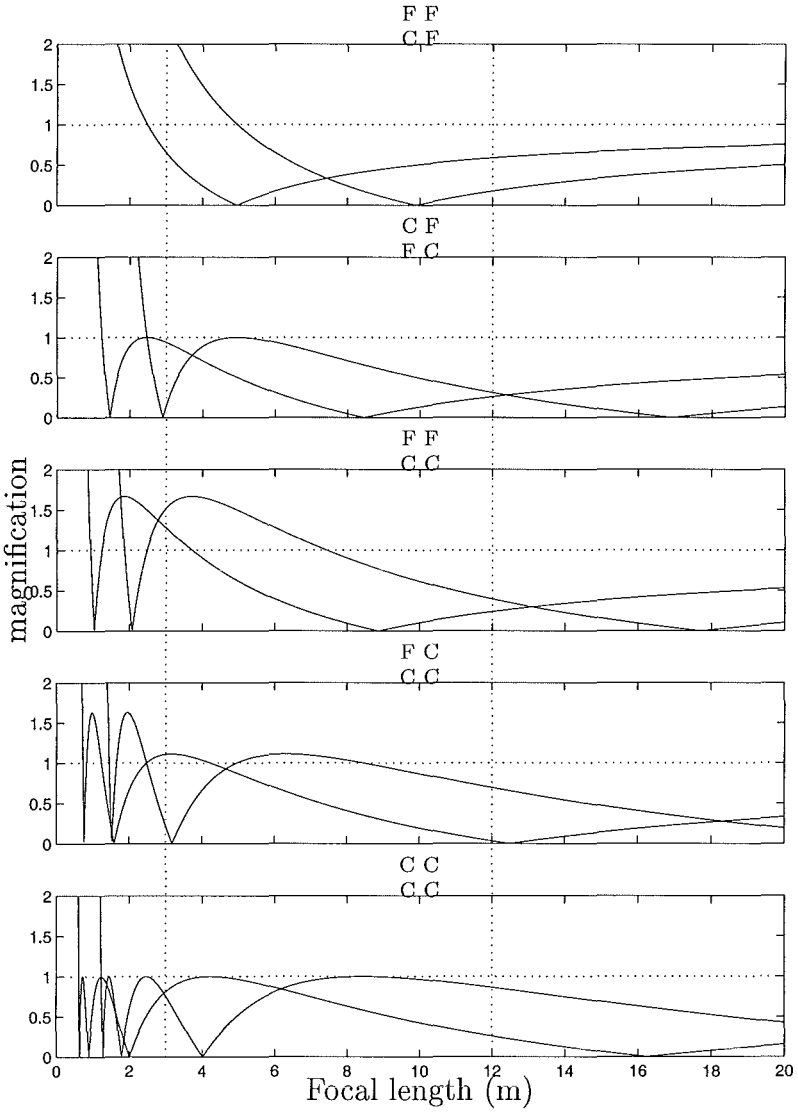


Fig. 6.1: The above magnification curves are used to determine cavity stability for the various mirror configurations. The focal lengths of particular interest here, 3 m and 12 m, are indicated on the plot. Each configuration has two magnification curves of the same functional form which relate to the in-plane (left curve) and out-of-plane (right curve) axes of the beam.

Stability is thus regarded here as the ability of a cavity to focus a plane wave. If a plane wave is not convergent after a circuit of the cavity then the cavity does not have the ability to contain a stable Gaussian mode. For the purposes of numerical calculations a plane wave is taken as a Gaussian beam which has a

spot size of 1 m at its waist. The cavity itself has been manufactured to within millimetre accuracy and is a square of side length 3 500 mm. The curved mirrors have a radius of curvature of 6 000 mm ($f = 3\,000$ mm). The in- and out-of-plane axes of the beam experience different curvature of the mirrors due to astigmatism (see Verdeyen [120, p.51]). The effective focal lengths experienced are found to be

$$f_{\text{out}} = f\cos\theta = \frac{3\,000}{\sqrt{2}}\,\text{mm}$$
$$f_{\text{in}} = \frac{f}{\cos\theta} = 3\,000\sqrt{2}\,\text{mm}\ .$$

Fig. 6.1 shows the magnification of the various cavity configurations as a function of mirror curvature with magnifications less than unity denoting stable cavities. Notice that each configuration has two curves of the same functional form which relate to the in-plane (left curve) and out-of-plane (right curve) axes of the beam. Of the two configurations which have two curved and two flat mirrors only the $\begin{smallmatrix}\text{C}&\text{F}\\\text{F}&\text{C}\end{smallmatrix}$ configuration is stable when f=3 m. Note that the proposed 12 m focal length mirrors (indicated on the plot) provide stable modes in whichever configuration they are used. This method of plotting cavity stability allows several interesting observations. For example if 5 m focal length mirrors are available then the cavity provides stable modes when only one curved mirror is used. Also if mirrors with focal length a little over 1 m are available then four may be used to provide stable lasing.

At the time that the initial work was performed on the large ring laser only two 6 m curvature mirrors were available. The $\begin{smallmatrix}\text{C}&\text{F}\\\text{F}&\text{C}\end{smallmatrix}$ arrangement is therefore the only configuration that provides stable lasing modes.

6.3 Ring Laser Design

G0 is a prototype for a slightly larger (P=16 m) high-cost device named the Grossring, or G, which will be installed in Wettzell, Germany. The current device was built to answer several key questions regarding the operation of the final model. These factors have had a large effect on the design of the laser cavity. In particular path length stability is not of great importance and hence the predominant material used here is stainless steel. This greatly reduces the total cost (in comparison with C-II and the proposed G) and has sped up the construction process due to availability of materials and ease of manufacture. The entire device was manufactured in the Department of Physics and Astronomy by Morrie Poulton and Rob Thirkettle.

In order to avoid problems of gas contamination, such as were experienced with C-I, no O-rings were used in G0. Instead all connections are made using all metal ConFlat™ flanges (CF275 and CF450). The corner boxes and gain section are mechanically isolated from the cavity pipes using bellows. In this way the

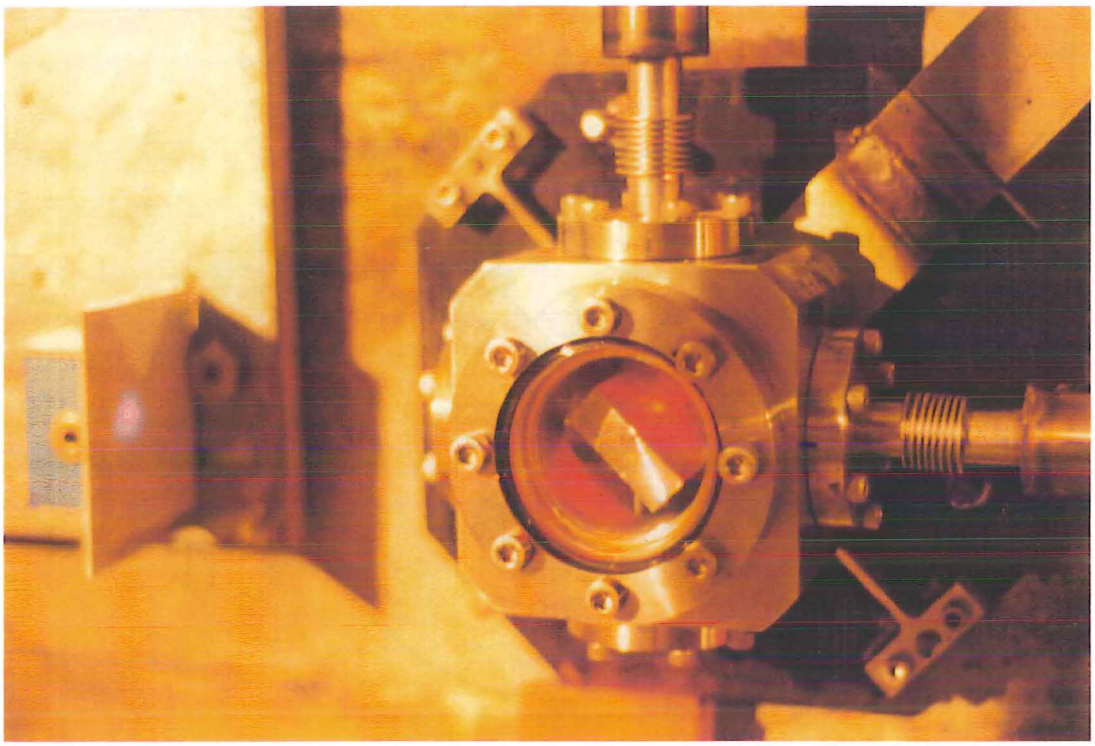


Fig. 6.2: This image shows the lower left (Eastern-Low) mirror housing. The mirror mount can be seen inside the chamber with the scattered light from the laser spot quite apparent. Note the blue halo around the exit beam seen on the screen mounted outside the laser cavity. This halo is due to the plasma that forms the gain medium.

pipes serve only to provide a vacuum manifold. Fig. 6.2 shows a picture of G0 shortly after it lased for the first time. Here the mirror holder, mirror and lasing spot can be seen through the CF450 window. This image should be compared with the line drawing given in Fig. 6.3 which provides a detailed layout of the mirror housing. Each mirror can be adjusted externally using a 1 m lever arm which is rigidly mounted to the mirror box. Fine adjustment of the lever is made using vernier screws which govern the mirror's tilt in-plane and out-of-plane. The mirror house is attached to the base plate (the dark octagonal plate in Fig. 6.2) via two H-sections which flex as the housing is adjusted.

The gain section is mounted so that the glass gain tube is mechanically isolated from the steel pipes of the manifold. Either end of the gain section may be adjusted using one of four vernier screws. Pumping is provided through a long bellows hose connected to the Eastern vertical pipe. A further CF133 connection is provided on the Western vertical arm in the event that analysis of the gas is required either using a ion pressure gauge or a mass spectrometer. The external optics are mounted on four optical tables mounted outside each of the mirror housings one of which is just visible in Fig. 6.2 with a magnetic clamp holding a screen in place. Because of the vertical arrangement the optical tables are made of steel sheet so that magnetic clamps can be easily used. These tables are also

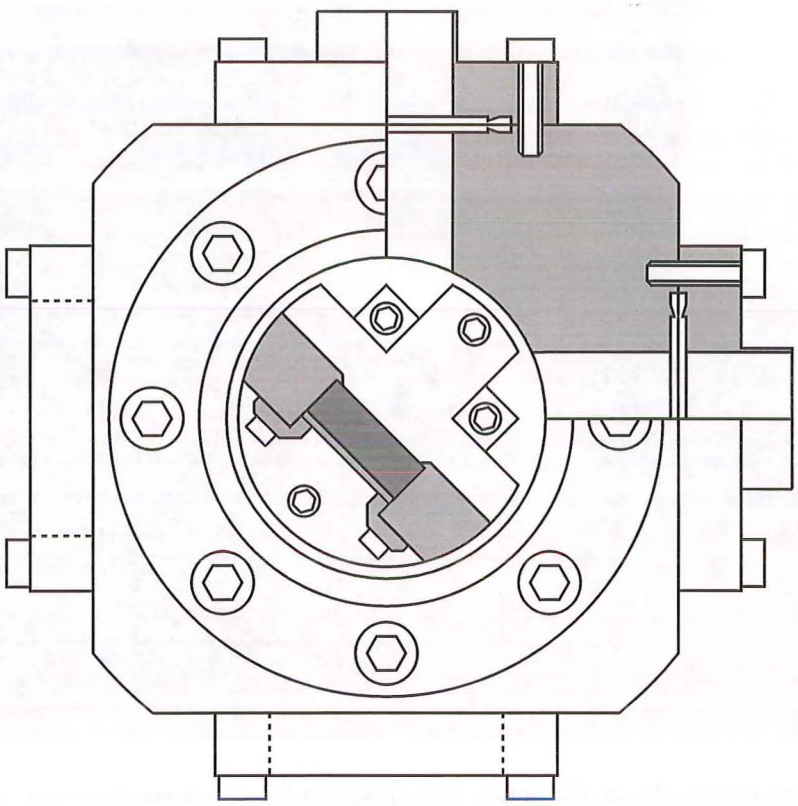


Fig. 6.3: This line diagram illustrates the construction of the mirror housings which are the key part of the laser design. Both the mirror holder and the top right quadrant of the box are shown in cross-section. This diagram should be viewed with reference to Fig. 6.2.

drilled and tapped on a 75 mm rectangular matrix.

6.4 Alignment Method

The alignment procedure is similar to that for C-I; a GreNe is used to align the cavity mirrors in much the same manner. There are several aspects of aligning such a large ring laser that are discussed here. Fig. 6.4 illustrates the layout of the optical components. Note that the ring is oriented vertically and has an arm length of 3.5 m. The alignment beam is injected into the ring cavity via three alignment mirrors, I1,I2 and I3. Typically only two mirrors are used for this but I1 and I2 were required as a beam ladder. Scattered light from the GreNe was visible on the front of M1 and this allowed the location to be easily adjusted to coincide with a desired area of the mirror. With such a long distance between mirror boxes it is difficult to inject the beam on-axis so that it emerges through the back of M2. To achieve this a taught fishing line is pulled between the two lower mirror housings. This is done with the line at the appropriate

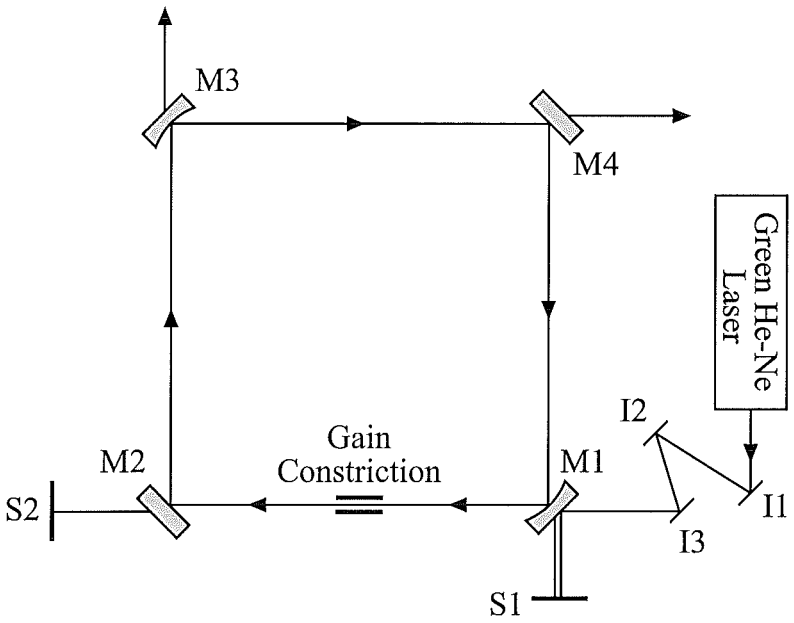


Fig. 6.4: This schematic illustrates the procedure which is used to align G0. The routine is straight-forward but the unusually large arm lengths require delicate adjustment for precise alignment.

height above the axis and allows an estimate of where the beam should strike I3. I3 is then adjusted to pass the beam on axis into the cavity. This trick is quite effective. Once achieved it is simple to walk the beam around (using I2 and I3) so that the spot on M1 remains stationary while the spot on M2 can be placed wherever desired. In order to minimise scatter effects the beam is required to pass axially through the gain constriction. This is achieved to high accuracy by adjusting either end of the constriction to the point where the beam grazes the glass tube. The distinctive pattern illustrated in Fig. 6.5 is then seen. When the beam strikes the interior of the tube below centre the reflected part is angled upwards as shown. The beam is raised until the *fan* pattern is parallel to the area vector of the ring laser. The beam is adjusted back through the middle of the tube until the same 'fan' pattern is seen on the other side. The optimal location is thus taken as the midpoint of these two extremes. The method is repeated for both ends of the gain tube. The circuit is completed by aligning M2, M3 and M4 in turn to place the beam on the desired part of the next mirror. With the return spot adjusted to coincide with the injected spot the final mirror, M1, is adjusted to make the alignment complete. Clearly using this process allows the beam to strike any part of any mirror. This is highly desirable in view of the method by which the mirrors are cleaned (see Chapter 2). The cleaning process leaves a particular part of the mirror, usually the centre, with a minimum of scatterers with the remainder of the mirror not necessarily as clean. These clean areas can be targeted by the alignment process and hence most primary scatterers can be avoided. The adjustment of M1, the final mirror, is achieved by viewing the exit beams on the screens S1 and S2. Note that the front- and back-reflected beams

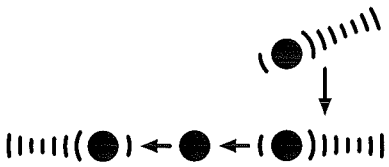


Fig. 6.5: This picture illustrates the sequence that has been adopted to pass the beam down the centre of the gain tube. The beam is scraped along the side of the tube to give the tail of interference fringes. This tail is adjusted to be horizontal. The symmetric pattern is found when the beam is scraped off the other side of the gain tube and the final position taken as the midpoint of the two extremes.

from M1 are seen when observed on S1 as indicated in Fig. 6.4. The circuit beam is required to align with the back reflected beam which is the left most on the diagram. Once M1 is roughly aligned the circuit beam is seen on S2 which may be located close to M2 or as far as 10 m away. The primary and circuit alignment beams are made coaxial by iteratively aligning the spots on S2 when in the near and far positions using M1 and M4 only. During these adjustments the double circuit beam is seen passing rapidly back and forth across S1. In this way the beams are made parallel to within 0.5 arc-second accuracy. At this stage fringing is visible which is due to interference between the primary and circuit beams. This constitutes the complete alignment process of the ring laser.

6.5 Radio Frequency Excited Gain Medium

Alterations have been made to the design of the r.f. gain coils. The system used to date on C-I has a primary coil of three or four turns. A further secondary coil is located to one side of the primary coil and comprises 12 to 16 turns. This has proved to be an efficient means of coupling the r.f. power into the discharge. However, an exploration carried out on C-I of the more simple design which uses only a primary coil has revealed that certain geometries will lase that have not in the past. In particular a three turn coil is placed on the Eastern arm of C-I which has an internal diameter of 12 mm. The increase of diameter of the gain tube naturally requires a drop in the operating pressure; the cavity is pumped from a 7:1 He:Ne mixture at 4 Torr to 2 Torr. In this configuration C-I lases requiring only 13 Watts of r.f. power and when pumped slowly continues to lase down to 0.4 Torr. The interesting feature of this method is the symmetry of the discharge which is complete as far as the eye can tell. This idea has been extended to G0; the secondary coil has been disconnected from the primary coil but is left wound around the glass tube. The discharge becomes very symmetric and is contained entirely within the gain capillary whereas with the old design the discharge extended well beyond the glass of the gain section and into the stainless steel tubes that form the laser cavity. The new discharge has the added benefit of being more efficient with only 5 Watts required to maintain lasing.

A fine tuning effect has been discovered entirely by accident. The secondary coil, now disconnected, has the ability to move the discharge along the gain capillary without significantly altering its profile. This provides added control over the precise location of the discharge which may prove useful in the future.

6.6 The Theoretical Earth Induced Sagnac Rate of G0

Here the expected output frequency due to the Earth's rotation is calculated. This is somewhat more complex than the equivalent calculation for C-I or C-II as the normal to the plane of the laser is not perpendicular to the Earth's surface. It is imagined that the wall on which G0 is mounted is located at the North pole. The scalar product between the unit vector pointing along the axis of the Earth's rotation $\hat{\Omega}$ and the unit vector pointing along the area vector of the ring $\hat{A}$ is required. To this end the unit area vector $\hat{A}$ located at the North pole is rotated through the latitude angle θ in such a way that after the rotation the vector points in the local North direction N (see Fig. 6.6). A further rotation

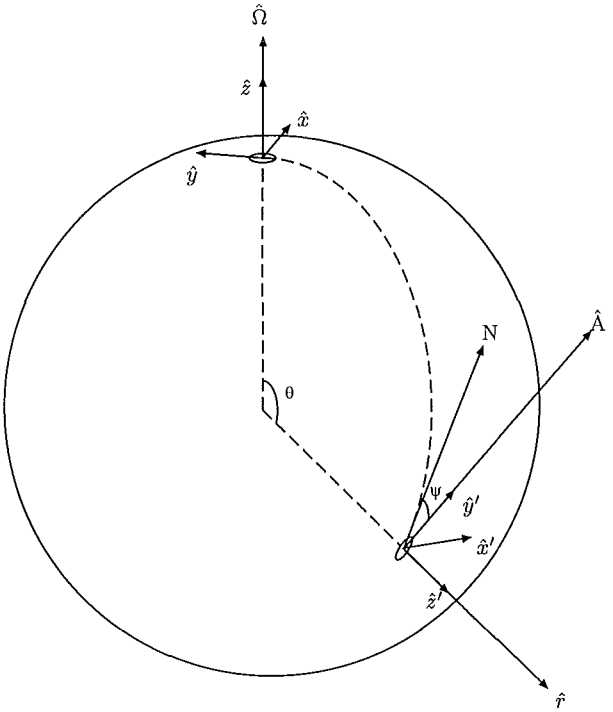


Fig. 6.6: The geometry used to find the scalar product of the unit area vector $\hat{A}$ and the unit Earth rotation vector $\hat{\Omega}$.

of $\hat{A}$ about the radial axis $\hat{r}$ through an angle ψ is then performed. The vector $\hat{A}$ is arbitrarily chosen to lie along the y-axis so that the first rotation is about the x-axis. The required scalar product is simply the direction cosine of $\hat{A}$ along

the z-axis after the two rotations. This scalar factor is found (using Geometric algebra [66, 67] but Euler angles could achieve the same result) to be $\sin\theta\cos\psi$. The theoretical Earth induced Sagnac rate is then given in terms of measurable parameters as

$$f_b = \frac{\Omega_E c}{4\lambda\nu_{\text{fsr}}}\sin\theta\cos\psi \ .$$

With the sidereal day taken as 86164 s the Earth's rotation rate is found. G0 has a free spectral range of 21.407 MHz from which a perimeter of 14.0044 m is deduced. The area can also be found from the perimeter measurement but the assumption that the ring is square is required which is reasonable in view of the fact that the area of the ring is independent of the misalignment angles (deviation from right-angles) to first order. With this assumption the area is $12.2577\,\text{m}^2$. The wavelength of the lasing light in the He-Ne gas mixture is taken as 633.0 nm and c is defined as $299792458\,\text{ms}^{-1}$ exactly. The latitude angle θ is 90° plus the cavern's latitude of $43°34'37''$ and the estimate of the angle ψ (to be improved using surveyor's records) is $10° \pm \frac{1}{2}°$. When inserted into the formula these values give a theoretical beat frequency of 287.75 ± 0.46 Hz.

6.7 G0 Output

Preliminary results are reported from G0. The cavity is filled with Helium and natural Neon in the ratio 6:1 up to a pressure of 3.5 mbar. Single longitudinal mode operation is achieved using exactly the same method as that used on C-I (see §2.8), that is, using the LDR/LED intensity servo-mechanism. While lasing in single longitudinal mode and unlocked under the Earth's rotation a Lissajous figure is generated on an oscilloscope with the Sagnac signal from G0 on the x-axis and a sinusoid (from a signal generator) on the y-axis. The figure can be made stationary by varying the frequency of the generated signal. Once achieved this becomes an extremely sensitive detector of phase shifts in the G0 output. A variety of tests can then be performed. Artificial seismic events are easily detected as a jitter in the Lissajous figure. Also the introduction of a magnetic field near the gain medium has the effect of shifting the beat frequency and hence causes the Lissajous figure to rotate at a constant rate. However, the question remains as to whether or not this signal is necessarily the Earth induced line. It is difficult to impose a rotation on the laser as it is rigidly mounted to a solid concrete wall which in turn is moulded around the solid volcanic rock that forms the interior cavern wall. It is conceivable that the signal can be traced back to some oscillation in the gain medium. This problem is yet to be clarified. However, it is assumed that the signal is indeed the Earth induced rate mostly because of the excellent agreement between the measured output frequency (see Figs. 6.7 and 6.8) and the expected frequency of 287.75 ± 0.46 Hz.

A series of data has been taken using the STROBES data acquisition software and is stored under the file name G29JA98B.SC2. The sample rate is 1000 samples

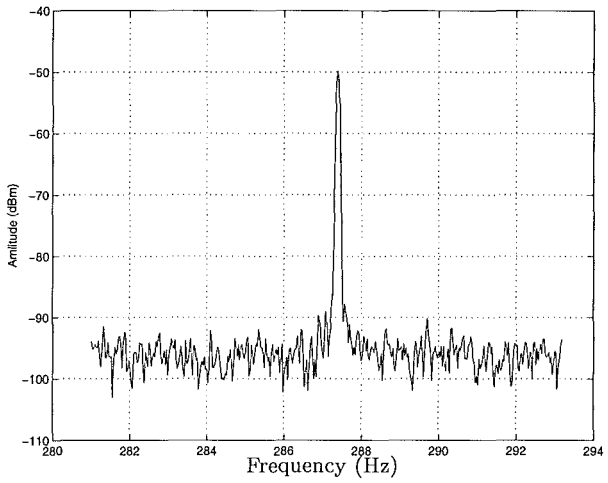


Fig. 6.7: Output from the audio spectrum analyser from Stanford Research Systems SRS770. This data was taken during the run G29JA98B.SC2 with 4 rms averages. The Earth induced Sagnac signal is well above the noise floor. Substantial effort was taken to keep the noise floor low. The cavern lights were turned off as was the vacuum equipment. The remaining noise is due to instrumentation (fans, electrical noise etc.) and background noise (micro-seismic, traffic etc.).

per second and the signal is passed through a 0-500 Hz anti-aliasing filter before A/D conversion. Unfortunately there is a great deal of polluting signal induced from mains power. This signal and its harmonics are removed in post-processing using a digital band-pass filter from 283-291 Hz. Frequencies are determined using the AR(2) technique discussed in §3.6. Fig. 6.8 displays the results of the analysis performed on the data set. The output frequency estimate is displayed as a function of time while the relative Allan deviation is determined directly from the frequency estimates. Notice that the minimum deviation occurs at about 3 s suggesting a highly unstable signal. The erratic behaviour at longer sample times (>40 s) is the result of there being only a small number of measurements for the longer sample times; the smaller number giving considerably larger uncertainties.

6.8 Conclusion

The operation of the G0 prototype has been a great success. Although the quantity of data available is small at the time of writing this thesis it provides the basis for several strong conclusions. The key conclusions to be drawn are that the design of this laser is sufficient to make such a large ring laser operate and that the servo mechanism is capable of holding the laser in single longitudinal mode, albeit only for tens of minutes. Further, the stability of the output frequency for the data set analysed is only ± 0.2 Hz drift in the carrier frequency of $\sim$287 Hz over the 15 minute measurement time. When the metal construction and the mounting on a concrete wall is considered it is remarkable that the carrier fre-

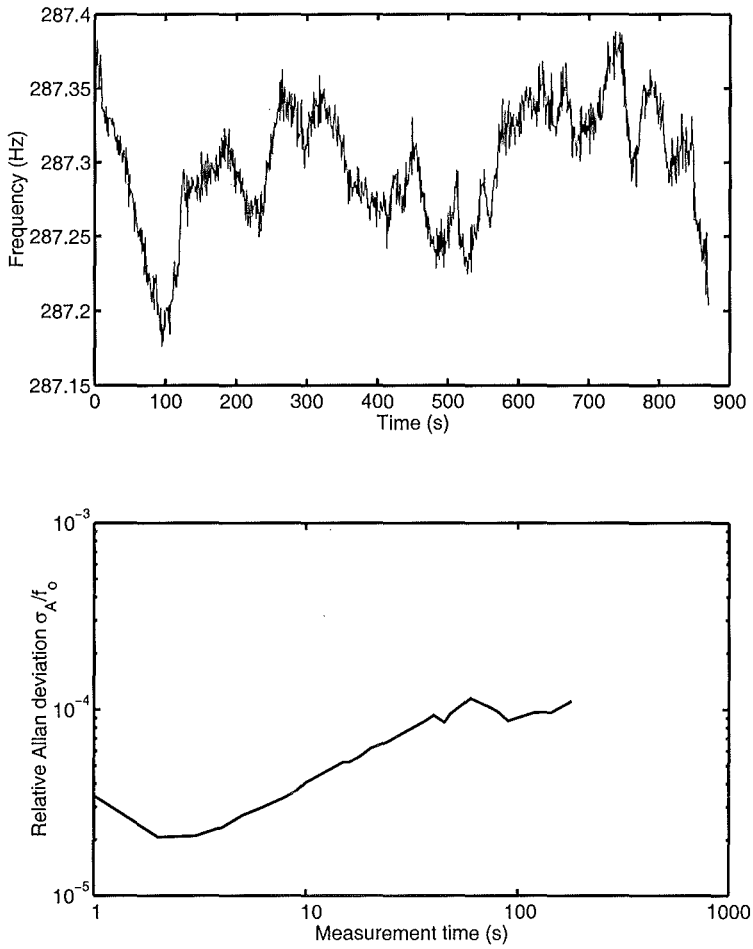


Fig. 6.8: This plot displays the carrier frequency from G0 as a function of time. The frequency estimates are made for every 1 s interval. The frequencies that are determined in the above method may be used to determine the Allan deviation of the signal. The Allan deviation has been calculated for different segment lengths and it is found that the segmentation of the data set has no effect on the Allan deviation.

quency should remain this stable. It is reasonable to conclude that the effects of coupling between the two modes have been greatly reduced by increasing the side length of the laser by a factor of 3.5. This goes some way towards verifying the quadratic scaling law for backscatter coupling discussed in the introduction. There remain several features of this large ring laser that have to be ascertained before a complete understanding of its behaviour is achieved. These include the classification of the laser's operating parameters such as the interdependence of output power with operating pressure, rf power etc. and more importantly the long term stability of the laser has to be determined. These matters are the subject of ongoing research within the ring laser group.

7. Conclusion

7.1 Summary of Results

In this thesis a wide variety of aspects of ring laser dynamics have been explored. All these diverse areas are part of a description of the deviation of real ring laser performance from the ideal. Several experimental achievements have been discussed. The decision to collect simultaneously the separate beam profiles has been pivotal and allowance for the observed differences between the beams has been a key part of the modelling process of the ring. A detailed and precise calibration procedure which is of particular application to this project has been developed. A considerable upgrade to the C-I ring laboratory has allowed these developments to proceed. In tandem with these data acquisition improvements the data reduction techniques have been advanced and allow the extraction of considerably more information than has been previously available. In particular the use of a second-order autoregressive model to describe the oscillating output has provided greater frequency resolution and accuracy. This method is also much faster computationally than its rivals.

A review of the fundamental limits placed on ring lasers has been provided. The quantum-noise from the discharge spontaneous emission is found to be the limiting factor in performance for short time scales. Environmental variations are identified as the major culprit accounting for frequency fluctuations for longer time scales. A clear presentation of these noise processes is obtained using the Allan deviation which allows identification of the noise types for different integration times. Identifying and reducing the dominant noise sources aids in the further reduction of the resolution of our ring lasers. The Allan deviation allows an identification of the quantum noise induced linewidth. A by-product of the second order autoregressive model provides a second measure of this linewidth and these two methods, which are entirely different in nature, agree well. Further, the linewidth indicates that our ring lasers operate close to the quantum limit for those integration times where quantum noise dominates.

A complex model has been developed which describes the oscillations of the irradiances of the two counter-propagating output beams. This model incorporates scattering phenomena that have been observed during routine operation of the laser. In particular the modulation of the scattered light is observed; the subsequent model is required to be consistent with this observation. Comparison between real data and numerical solutions of the model show that frequency

shifts and waveform distortion are accurately described by the model. These frequency shifts may be either side of the theoretical beat frequency by several hertz; the model successfully provides a close fit waveform in either case. Other tests of the model include two locking phenomena. The first locks the beat frequency to an externally applied modulation of the perimeter. The nature of this locking is determined experimentally and the predictions of the model are qualitatively in good agreement with these measurements. The second test synthesises the behaviour of the model as the rotation rate is varied. Although no measurements of this are currently possible the predictions compare well with the literature [45]. When locked to an external modulation many averages of the audio spectrum reveal low amplitude effects which is only possible due to the absence of carrier drift. While in this operating mode small, spectrally broad modulations of the carrier are revealed between 0.1 and 0.3 Hz. While an explanation of these modulations is not provided the potential for a new, highly sensitive mode of operation is demonstrated

The final result of the work reported here is the commissioning of a prototype very-large (14 m perimeter) ring laser. While the design and manufacture of this prototype were the results of efforts from the whole group I provide a detailed account for completeness.

7.2 Further Work

Some of the ongoing work which is discussed here is already (at the time of completion of this thesis) under way. Other aspects may fit into the ongoing ring laser programme. Most of the material mentioned here is equally applicable to any ring laser gyro existing or proposed.

Certain experimental details have not been addressed or have not be followed to a satisfactory conclusion in the time I have had available. The intensity servo mechanism works well for short times but invariably drifts over a day or so. A sophisticated servo-mechanism requires considerable control theory even for a problem as simple as ours. Also, the calibration procedure is laborious and automation would save a lot of experimenting time. With output power stabilised for long periods of time, gas contamination would become of greater importance. The gas bleed-through system, to date only operated manually, would have to become automated. Another separate servo-mechanism would control a fine leak valve to keep the pressure in the lasing cavity constant. These measures could increase the operating duration from days to months.

The process of identifying and reducing the dominant noise source is ongoing. Ultimately flicker noise will prevent the reduction of the overall stability. Flicker noise is independent of integration time and hence presents a basic limit to performance. A complete study of flicker noise in ring lasers would be an enormous undertaking. None of the analyses reported by this group are limited by flicker

noise but as stability improves this noise will certainly arise. The agreement between the predictions of the Allan deviation and the second order autoregressive model is good. Qualitatively it would appear that the correspondence between the two methods could be improved. A greater sophistication in the model, perhaps by introducing a moving average, could improve the agreement.

The story of quantum limits in ring lasers is not yet complete. The nature of the phase difference between two modes which share the same cavity but which are degenerate may be explored using the phase experiment proposed in §4.3.5. This experiment may provide vital information about ring lasers which have extremely low light levels and hence require a more sophisticated description of phase.

The characterisation of the model described in Chapter 5 is incomplete due to limitations imposed by computer resources. Ultimately this model could be used to correct the output beat frequency. Waveform distortions and phase shifts between counter-propagating beams contain information about push/pulling effects which may be used to correct the output. The potential benefits are so great that this procedure is already being pursued.

Appendices

A. The Hartley Transform

The main theme of Chapter 3 is to consider methods of analysis other than the Fourier transform. As an aside I mention here an alternative formulation of spectral analysis which is due to Hartley [56] and is discussed in detail in a book by Bracewell [23]. The key points regarding this transform are provided here so that a concrete example is available. The transform pair is

$$H(f) = \int_{-\infty}^{\infty} V(t) \operatorname{cas} 2\pi f t dt \tag{A.1}$$

$$V(t) = \int_{-\infty}^{\infty} H(f) \operatorname{cas} 2\pi f t dt \tag{A.2}$$

where $V(t)$ is the original waveform, $H(f)$ is the Hartley transform and the cas function is defined by

$$\operatorname{cas} t \equiv \cos t + \sin t \ . \tag{A.3}$$

The discrete form of the transform is formulated in the same manner as the discrete Fourier transform. There are certain features of the transform that are worth mentioning. $H(f)$ is real valued which is often considered a desirable feature as the power spectrum is itself real. All information contained in the Fourier transformed data is also contained in the Hartley transformed data; for example the power spectrum is given by

$$S(f) = H^2(f) + H^2(-f) \tag{A.4}$$

and the phase of the complex spectral component is given by

$$\phi(f) = \tan^{-1}\left(\frac{H(-f) - H(f)}{H(f) + H(-f)}\right) \ . \tag{A.5}$$

Further, if the function $V(t)$ is a purely odd or purely even function then the Hartley transform is equivalent to the Fourier transform.

The Hartley transform has a fast algorithm in the same vein as the Fast Fourier Transform but which makes use of the fact that the starting values are real and hence achieves a factor 2 speed improvement over the Fast Fourier Transform of a data set of the same length. However more recent implementations of the Fast Fourier Transform have managed to equal this factor of two improvement (see § 3.2.2).

B. Averaging Proof

This appendix contains a proof that the probability distribution of the rms average of n numbers chosen from $P_1(r) = 2kre^{-kr^2}$ where $r > 0$ is given by

$$P_n(r) = \frac{2(nk)^n}{(n-1)!} r^{2n-1} e^{-nkr^2} \ . \tag{B.1}$$

Proof The proof is by induction. Clearly the result holds for $n = 1$. Assuming the result holds for n, $P_{n+1}(r)$ is found

$$P_{n+1}(r) = \frac{\int_0^{\sqrt{\frac{n+1}{n}}} P_n(s) P_1(\sqrt{(n+1)r^2 - ns^2}) \mathrm{d}s}{\int_0^{\infty} \left[\int_0^{\sqrt{\frac{n+1}{n}}} P_n(s) P_1(\sqrt{(n+1)r^2 - ns^2}) \mathrm{d}s \right] \mathrm{d}r} \ . \tag{B.2}$$

The top line is evaluated by inserting the $P_n(r)$ and $P_1(r)$ expressions and rationalising

$$\frac{2(nk)^n 2k}{(n-1)!} e^{-(n+1)kr^2} \int_0^{\sqrt{\frac{n+1}{n}}} s^{2n-1} \sqrt{(n+1)r^2 - ns^2} \mathrm{d}s \ . \tag{B.3}$$

Making use of the identity

$$\int_0^{\sqrt{\frac{n+1}{n}}} s^{2n-1} \sqrt{(n+1)r^2 - ns^2} \mathrm{d}s = \frac{(n+1)^{n+\frac{1}{2}} (n!)^2 2^{2n-1}}{n^{n+1}(2n+1)!} r^{2n+1} \tag{B.4}$$

so that the top line is

$$\frac{2(nk)^n 2k}{(n-1)!} \frac{(n+1)^{n+\frac{1}{2}} (n!)^2 2^{2n-1}}{n^{n+1}(2n+1)!} e^{-(n+1)kr^2} r^{2n+1} \ . \tag{B.5}$$

From this the bottom line is then evaluated by use of standard integrals to infinity of Gaussian functions

$$\frac{2(nk)^n 2k}{(n-1)!} \alpha(n) \frac{n!}{2((n+1)k)^{n+1}} \tag{B.6}$$

from which it is found that

$$P_{n+1}(r) = \frac{2((n+1)k)^{n+1}}{n!} r^{2n+1} e^{-(n+1)kr^2} \ . \tag{B.7}$$

This is the result for $n+1$ and thus the result holds for all n by induction. ■

C. Neon Fact Sheet

A variety of information relating to neon has been required throughout this thesis. Neon is of course the source of gain in the He-Ne laser and hence its properties are of utmost importance. Neon was discovered by Sir William Ramsey and M. W. Travers in 1898 and its name is taken from the Greek *neos* meaning new. Neon forms (18.18±0.04) ppm of air.

Isotope	Abundance(%)	Atomic Mass	Spin	Decay modes
$_{10}Ne^{20}$	90.51	19.992 44 AMU	0	stable
$_{10}Ne^{21}$	0.27	20.993 95 AMU	3/2	stable
$_{10}Ne^{22}$	9.22	21.991 38 AMU	0	stable

Table C.1: This table provides key nuclear information about the naturally occuring neon isotopes. Note that the definition of atomic mass assigns the value 12.00000 to $_6C^{12}$ which results in one AMU = $1.660\,565\,5 \times 10^{-27}$ kg (±5.1 ppm). From CRC Handbook [124].

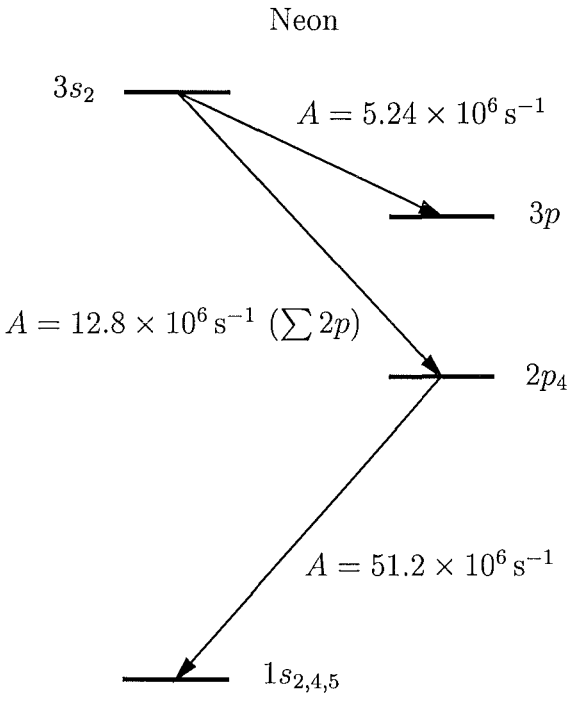


Fig. C.1: The primary decay paths for the upper and lower levels of the 632.8 nm transition in neon. Data is taken from Verdeyen [120].

The λ = 632.82 nm (633.00 nm in vacuum) emission line is due to the transi-

tion ($3s_2 \rightarrow 2p_4$) where the upper energy level is 166 657 cm^{-1} above the ground state. The upper and lower levels of this transition have the degeneracies 3 and 5 respectively. The transition probability (Einstein A coefficient) is given in the CRC Handbook [124] as $0.039 \times 10^8 s^{-1}$. A table of thermal properties is given in Table C.2

Property	Value	Unit
Melting point (at 1 atmosphere)	-248.67	°C
Boiling point (at 1 atmosphere)	-246.048	°C
Specifid heat	0.904	$J\,g^{-1}\,K^{-1}$
Heat of fusion	0.3317	$k\,J\,mol^{-1}$
Heat of vapourisation	1.7326	$W\,cm^{-1}\,K^{-1}$
Density	0.0008999	$g\,cm^{-3}$
1^{st} ionisation potential	21.5645	eV
2^{nd} ionisation potential	40.962	eV
3^{rd} ionisation potential	63.45	eV

Table C.2: This table provides a list of important thermal and ionisation properties for neon.

The mean free path is given by Harris [55] $6.4 \times 10^{-3}/P$ cm where the pressure P is required in m Bar. The rms velocity of atoms in the gas is given by

$$v_{\text{rms}} = \sqrt{\frac{3k_B T}{m}} \tag{C.1}$$

where k_B is Boltzmann's constant equal to 1.38×10^{-23} J molecule^{-1} K^{-1}. Taking the temperature T to be approximately 293 K and the molecular mass of Ne^{20} as 3.35×10^{-26} kg molecule^{-1} it is found that the root mean square velocity of neon at room temperature is $\approx 600\,m\,s^{-1}$.

D. Random Walks

D.1 Random Walks in One and Two Dimensions

Random walks are of importance in quantum noise calculations. Spontaneous emission into a lasing mode makes the phase of the mode diffuse away from the ideal linear phase evolution. This phase diffusion constitutes a one-dimensional walk in phase space. The thorough analysis of random walks due to Chandrasekhar [26] (found in the collection of papers on related topics by Wax [123]) is summarised here. Some examples relevant to coherent light fields are worked through starting with the initial phase diffusion due to spontaneous emission. Other examples refer to walks in two dimensions and are relevant to the amplitude-phase evolution of a spontaneous emission field. Further discussion of the role that walks play in describing the quantised radiation field is provided by Loudon [72, Chapter 3,p.103]

D.1.1 Chandrasekhar's Approach

Chandrasekhar's work involves a very thorough analysis of walks or flights as he calls them. He derives a general mechanism for finding the probability distribution for the length of a walk of N steps in D dimensions in which each step is governed by a probability distribution $\tau(\boldsymbol{r})$. Chandrasekhar's emphasis is on the use of such a scheme for determining the behaviour of the gravitational field in a random distribution of stars (a star cluster). He is interested in fluctuations which occur as a matter of course and which may explain some of the curious effects which are seen; such as a star which is suddenly (on cosmic time scales) ejected from a star cluster. For these reasons all of his worked examples pertain to walks of vectors which belong to a particular distribution and which are contained within flat 3-dimensional space.

Consider a walk, $\boldsymbol{R} = \sum_{j=1}^{N} \boldsymbol{r}_j$, in D-dimensional space which is made up of N steps where each step is found from a probability distribution $\tau_j(q_j^1, \ldots, q_j^s)$ which is the probability that the j^{th} step has length $\boldsymbol{r}_j(q_j^1, \ldots, q_j^s)$ and where $j = 1 \ldots N$. The probability that the final vector lies inside a volume $\boldsymbol{R}$ to $\boldsymbol{R} + d\boldsymbol{R}$ is given

by

$$W_N(\boldsymbol{R}) = \frac{1}{(2\pi)^D} \int \cdots \int \exp(-i\boldsymbol{\rho} \cdot \boldsymbol{R}) A_N(\boldsymbol{\rho}) d\boldsymbol{\rho} \tag{D.1}$$

where $A_N(\boldsymbol{\rho})$ is written

$$A_N(\boldsymbol{\rho}) = \prod_{j=1}^{N} \int \cdots \int \tau_j(q_j^1, \ldots, q_j^s) \exp(i\boldsymbol{\rho} \cdot \boldsymbol{r}_j(\boldsymbol{q})) dq_j^1 \ldots dq_j^s \ . \tag{D.2}$$

The case of interest is that when all the distributions τ_j are equal in which case the last expression reduces to:

$$A_N(\boldsymbol{\rho}) = \left(\int \exp(i\boldsymbol{\rho} \cdot \boldsymbol{r}(\boldsymbol{q})) \tau(\boldsymbol{q}) d\boldsymbol{q} \right)^N \ . \tag{D.3}$$

These equations summarise that part of Chandrasekhar's work which is relevent to the phase diffusion caused by quantum noise. The process of finding the final probability distribution for the length of the walk, $W_N(\boldsymbol{R})$, requires that the Fourier Transform of the probability distribution, $\tau(\boldsymbol{q})$, of each step be found, this function raised to the N^{th} power and the inverse Fourier Transform found. It is during the intermediate step of raising the transformed distribution to the N^{th} power that the limits of large N are imposed.

D.1.2 One Dimensional Walks

Phase is confined to the unit circle S^1. The assumption is made here that the phase contributions are sufficiently small that, during time intervals much shorter than the time between samples, the quantum noise phase contribution has not shifted the cavity phase through a significant fraction of a full cycle. This assumption is quite satisfactory since the millihertz linewidths found at the end of Chapter 4 require $\approx 1000\,\text{s}$ to diffuse through a full cycle whereas the samples are usually taken 1 ms apart. The example considered here is that of a walk with Gaussian distributed lengths so that $\tau = \sqrt{k/\pi} e^{-kr^2}$ which provides the expression for $A_N(\rho)$

$$\begin{aligned} A_N(\rho) &= \left[\int_{-\infty}^{\infty} \sqrt{k/\pi} e^{-kr^2} e^{i\rho r} dr \right]^N \\ &= \exp\left(-\frac{N\rho^2}{4k} \right) \end{aligned} \tag{D.4}$$

Inserting this into the expression for $W_N(R)$

$$\begin{aligned} W_N(R) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp\left(-\frac{N\rho^2}{4k} \right) \exp(-i\rho R) d\rho \\ &= \sqrt{\frac{k}{N\pi}} \exp(-\frac{kR^2}{N}) \ . \end{aligned} \tag{D.5}$$

As expected the probability distribution of the phase spreads as the number of steps (or time) increases. Chandrasekhar also shows how under the appropriate conditions the probability density $W_N(\boldsymbol{R})$ evolves according the familiar diffusion differential equation

$$\frac{\partial W_N(\boldsymbol{R})}{\partial t} = D\nabla^2 W_N(\boldsymbol{R}) \tag{D.6}$$

and a complete analysis of phase diffusion would require the solution of this differential equation around an annulus with periodic boundary conditions. As stated above this is not necessary in this situation.

D.1.3 Walks in Two Dimensional Space

A further worked example of Chandrasekhar's method is provide here. This example comes from the consideration of the amplitude and phase of a chaotic light beam. Each walk comprises unit steps corresponding to the contribution due to each photon. Loudon [72, Chapter 3,p.103] refers to the result: for a random walk of N unit steps in 2-dimensions the probability that the end of the walk is inside a small area $r\,dr\,d\theta$ located at (r,θ) is

$$\frac{1}{\pi N}\exp(-r^2/N)\ . \tag{D.7}$$

Here I rederive this result using Chandrasekhar's generalised method. The probability distribution for each step of the random walk is $\tau(\boldsymbol{r}) = \delta(r^2-1)/2\pi$. Inserting this into the equation for $A_N(\boldsymbol{\rho})$

$$\begin{aligned} A_N(\boldsymbol{\rho}) &= \left[\frac{1}{2\pi}\int_{-\pi}^{\pi}\int_0^{\infty}\exp(i\rho r\cos\theta)\delta(r^2-1)r\,dr\,d\theta\right]^N \\ &= \left[\frac{1}{2\pi}\int_{-\pi}^{\pi}\exp(i\rho r\cos\theta)d\theta\right]^N \\ &= \left[\frac{1}{2\pi}\int_{-\pi}^{\pi}\cos(\rho r\cos\theta)d\theta\right]^N \end{aligned} \tag{D.8}$$

where the integrals over the infinite two dimensional surface are presented in polar coordinates and θ is the angle between $\boldsymbol{r}$ and $\boldsymbol{\rho}$. Expanding the inner cosine term and then performing the integral to get a series expression for the result:

$$\begin{aligned} A_N(\boldsymbol{\rho}) &= \left[\frac{1}{2\pi}\int_{-\pi}^{\pi}\left\{1+\sum_{n=1}^{\infty}\frac{(-1)^n}{(2n)!}(\rho\cos\theta)^{2n}\right\}d\theta\right]^N \\ &= \left[1+\frac{1}{2\pi}\sum_{n=1}^{\infty}\frac{(-1)^n}{(2n)!}\rho^{2n}\int_{-\pi}^{\pi}\cos^{2n}d\theta\right]^N \\ &= \left[1+\sum_{n=1}^{\infty}\frac{(-1)^n}{(n!)^2}\left(\frac{\rho}{2}\right)^{2n}\right]^N\ . \end{aligned} \tag{D.9}$$

Now if $p(x)$ is a polynomial in x where a_k is the first non-zero coefficient so that $p(x) = a_k x^k + a_{k+1}x^{k+1} + \ldots$ then in general

$$\lim_{N\to\infty} [1 + p(x)]^N = \exp(N a_k x^k) \tag{D.10}$$

which is independent of the coefficients a_{k+1}, a_{k+2} etc. This result allows the great simplification

$$\lim_{N\to\infty} A_N(\boldsymbol{\rho}) = \lim_{N\to\infty} \left[1 + \sum_{n=1}^{\infty} \frac{(-1)^n}{(n!)^2} \left(\frac{\rho}{2}\right)^{2n}\right]^N = \exp(-\frac{N\rho^2}{4}) \; . \tag{D.11}$$

Next this simplified form is substituted into the expression for $W_N(\boldsymbol{R})$:

$$W_N(\boldsymbol{R}) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \exp(-i\boldsymbol{\rho}.\boldsymbol{R}) \exp(-\frac{N\rho^2}{4}) d\boldsymbol{\rho} \; . \tag{D.12}$$

This integral is expressed in terms of polar variables (r, θ)

$$W_N(\boldsymbol{R}) = \frac{1}{4\pi^2} \int_{\mathbb{E}^2} \exp(-i\rho R\cos(\theta)) \exp(-\frac{N\rho^2}{4}) \rho d\rho d\theta \tag{D.13}$$

which is evaluated with the help of MATHEMATICA®

$$W_N(\boldsymbol{R}) = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \frac{\sqrt{\pi}\left(-2iR\cos\theta + \sqrt{N\pi}\ \text{Laguerre}(\frac{1}{2}, -\frac{1}{2}, x)\right)}{N^{3/2}\exp(x)} d\theta \tag{D.14}$$

where $x = R^2\cos^2\theta/N$. This expression is simplified by noticing that the imaginary part vanishes due to symmetry of the cosine terms on the interval $(-\pi, \pi)$. Next the Laguerre polynomial is expanded according to

$$\text{Laguerre}(\tfrac{1}{2}, -\tfrac{1}{2}, x) = -\frac{2}{\pi} \sum_{n=0}^{\infty} \frac{x^n}{n!(2n-1)} \tag{D.15}$$

and the exponential is also expanded and the product of the sequences is raationalised so that:

$$\begin{aligned}
W_N(\boldsymbol{R}) &= \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \frac{\pi}{N}\frac{2}{\pi} \left(\sum_{n=0}^{\infty} \frac{-x^n}{n!(2n-1)}\right)\left(\sum_{n=0}^{\infty} \frac{(-x)^n}{n!}\right) d\theta \\
W_N(\boldsymbol{R}) &= \frac{1}{4\pi^2}\frac{2}{N} \int_{-\pi}^{\pi} \left(\sum_{n=0}^{\infty} \frac{(-x)^n 2^{2n} n!}{(2n)!}\right) d\theta \\
W_N(\boldsymbol{R}) &= \frac{1}{4\pi^2}\frac{2}{N} \left(\sum_{n=0}^{\infty} \left(\frac{-R^2}{N}\right)^n \frac{2^{2n} n!}{(2n)!} \int_{-\pi}^{\pi} \cos^{2n}\theta d\theta\right) \; .
\end{aligned} \tag{D.16}$$

Making use of the identity

$$\int_{-\pi}^{\pi} \cos^{2n}\theta d\theta = \frac{(2n)!\pi}{(n!)^2 2^{2n-1}} \tag{D.17}$$

the result reduces to

$$\begin{aligned} W_N(\boldsymbol{R}) &= \frac{1}{4\pi^2}\frac{2}{N}\sum_{n=0}^{\infty}\left(\frac{-R^2}{N}\right)^n \frac{2^{2n}n!}{(2n)!}\frac{(2n)!\pi}{(n!)^2 2^{2n-1}} \\ W_N(\boldsymbol{R}) &= \frac{1}{N\pi}\exp(\frac{-R^2}{N}) \ . \end{aligned} \tag{D.18}$$

This is the result that Loudon refers to. Returning to the original problem, the expectation value of the length of a walk of N unit steps is required. This is found by using the above formula in cylindrical polars:

$$\begin{aligned} \langle R\rangle &= \int_{\mathbb{E}^2} RW_N(\boldsymbol{R})dA \\ &= \int_{-\pi}^{\pi}\int_0^{\infty}\frac{1}{N\pi}\exp(\frac{-R^2}{N})R^2 dRd\theta \\ &= \frac{\sqrt{N\pi}}{2} \end{aligned} \tag{D.19}$$

and as expected the expectation value of the length of a walk in two dimension increases as the square root of the number of steps in the walk but with a proportionality constant of $\sqrt{\pi}/2$.

E. Previous Drift Correction Analyses

Two closely related methods of determining the real Sagnac rate from a pulled signal are discussed here. These methods are discussed in Stedman et. al. [112] although the reanalysis shown here does not appear to be globally repeatable. In its essence the technique uses the fact that if a ring laser's output is pulled it will also be distorted in a particular way. We determine the parameters which define the amount of pulling; from them a correction factor is determined. We seek a relationship of the form $f_s = kf_b$ where f_s is the ideal Sagnac frequency (at $\sim 68.95\,\mathrm{Hz}$ for CI and $\sim 79.40\,\mathrm{Hz}$ for CII), f_b is the actual frequency of the output and k is the correction factor which is yet to be determined. Two different methods are discussed here each of which provides a correction factor k. For each of these methods experimental results are included which illustrate the correction process. Both methods rely on the Adler equation to describe the pulling that arises from mode coupling:

$$\dot{\psi}(t) = f_s + f_l \sin(\psi(t)) \tag{E.1}$$

which has a solution for $f_s > f_l$ that is well known and is found explicitly in § 5.1.2

$$\psi(t) = 2\tan^{-1}\left(\frac{f_l + f_b \tan(\pi f_b t)}{f_s}\right) . \tag{E.2}$$

Two data sets have been chosen for this illustration, one from each of CI and CII. Both are taken using a photomultiplier tube as detection device. The data from CI has an unusually stable output frequency but represents single beam output. We should not expect too much of the results obtained from this data as the model used here applies only to a combined beam interferogram. However, it has been noted that the harmonic content of single and combined beam outputs are very similar and hence the analysis may give useful indications. The second data set is taken from CII and represents the most frequency stable output to date (July 1997). This output arises from combined beams and hence is more relevant to the analyses in question than the data set from CI. This CII data also illustrates the greatest flaw in the analyses; the output frequency is greater than the theoretically expected value of 79.40 Hz. Both theoretical models assume Adler pulling which has no mechanism for frequency pushing.

E.1 Instantaneous Frequency Method

The first method uses the harmonic structure of the instantaneous frequency to determine the correction to the output beat frequency f_b. These calculations follow Ziyuan Li [128]. The technique requires a Fourier analysis of the instantaneous frequency

$$f_i = \frac{1}{2\pi}\frac{d\Psi(t)}{dt} = \frac{f_b^2 f_s}{f_l(f_l\cos(2\alpha) + f_b\sin(2\alpha)) + f_l^2 + f_b^2} \tag{E.3}$$

where we write $\alpha = \pi f_b t$. Now we note that:

$$f_l\cos(2\alpha) + f_b\sin(2\alpha) = f_s\sin(2\alpha + \tan^{-1}(f_l/f_b)) \tag{E.4}$$

and write $\epsilon = \tan^{-1}(f_l/f_b)$ so that after simplification

$$f_i = \frac{f_b^2}{f_s + f_l\sin(2\alpha + \epsilon)} \,. \tag{E.5}$$

We choose the origin of time so that $\epsilon = \pi/2$ and the resulting function is even. The Fourier components are given by

$$F_n = \frac{2}{\tau}\int_0^\tau \cos(2\pi nt/\tau)\frac{f_b^2}{f_s - f_l\cos(2\pi f_b t)}dt \,. \tag{E.6}$$

Writing $\phi = 2\pi t/\tau$ we find

$$\begin{aligned} F_n &= \frac{f_b^2}{\pi f_s}\int_0^{2\pi}\frac{\cos(n\phi)}{1 - \frac{f_l}{f_s}\cos\phi}d\phi \\ &= \frac{f_b^2}{\pi f_s}\int_0^{2\pi}\frac{e^{in\phi}}{1 - \frac{f_l}{f_s}\cos\phi}d\phi \quad \text{noting that} \quad \int_0^{2\pi}\frac{\sin(n\phi)}{1 - \frac{f_l}{f_s}\cos\phi}d\phi = 0 \end{aligned} \tag{E.7}$$

This integral may be evaluated by using the substitution $z = e^{i\phi}$ and integrating around the unit circle. This involves some detail and application of Cauchy's Residue Theorem and hence only the result is quoted:

$$F_n = 2f_b\left(\frac{f_s - f_b}{f_l}\right)^n \,. \tag{E.8}$$

The harmonics are thus in geometric progression which means that on a logarithmic plot they will form a linear plot. The scale factor of the progression is $r = (f_s - f_b)/f_l$. We are now in a position to eliminate f_l using the standard result of Adler locking theory $f_s^2 = f_l^2 + f_b^2$. The solution is simply found:

$$f_s = f_b\left(\frac{r^2+1}{r^2-1}\right) \,. \tag{E.9}$$

E.1.1 A Note on Data Reduction

The process of finding the heights of the Fourier components requires a correction to be made to the spectrum of the instantaneous frequency. When finding the numerical instantaneous frequency we first find the phase from the analytic signal. Next the phase is differentiated by use of a discrete time differential filter; in this case $[-1, 1]$ is used. The frequency response of this filter has to be accounted for. We find the transfer function of the filter and correct the spectrum accordingly. The transfer function is simply given by $H(\omega) = \sqrt{2 - 2\cos\omega}$. The correction made here is necessary for any attempt at determining f_s using this procedure. The corrected spectrum is shown in Fig.(E.1). The data sets

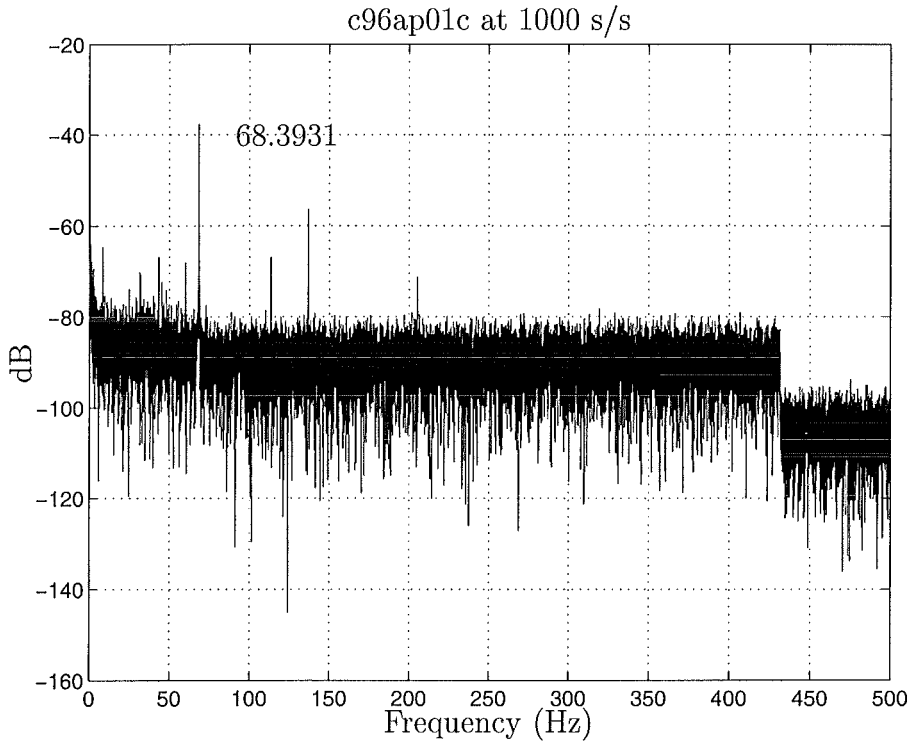


Fig. E.1: The CI data set used is the run from 1 April 96 named c96ap01c.scl from which the first 2^{15} samples are taken.

c96ap01a.scl and c96ap01b.scl are not used as it was not certain that these were taken in the absence of an FSR carrier at 86.410 MHz. These data sets contain Sagnac frequencies in excess of the theoretical Sagnac rate suggesting the presence of more than a single mode. It is certain that the data set used here was taken in the absence of a FSR carrier. Most importantly the data is taken from a single beam. The theory developed here regards interferometrically combined beams. The hope is that the theories will be strongly related but this is yet to be verified. The unwrapping procedure gives rise to the step in the noise spectrum at f_b below the Nyquist frequency. The spectrum shown has been corrected for the frequency response of the digital differentiation.

E.1.2 Results From Instantaneous Frequency Method

We first establish an accurate value for the current beat frequency. We achieve this by noting that the average frequency over a period is equal to f_b. We can therefore take the average of all the instantaneous frequencies to get an accurate value for f_b. Note that the length of a run may not correspond to an integer number of periods, making the above statement invalid. If the sum is done over many periods then any error due to this effect will be divided out and made insignificant. With the data set c96ap01c.sc1 we have just over 2241 periods giving an over estimate of the uncertainty as ± 0.003 Hz. We find the above average to be (68.393 ± 0.003) Hz.

Upon analysis the carrier frequency has a spectral line with height -37.741 dB and first and second harmonics with heights of -56.447 dB and -71.336 dB respectively. The length of the data set was chosen in such a manner that the carrier (and hence its harmonics) fall fortuitously almost exactly on a frequency bin. This method minimises the amount of uncertainty in the measurement of the height of the lines. An alternative procedure, not implemented, uses the power content of the spectral lines summed over a few neighbouring frequency bin. The differences in the heights of the lines is 18.706 dB and 14.889 dB. From these we calculate the scaling factor, for example, for the first ratio

$$r = 10^{18.706/20} = 8.616\ldots .$$

This gives a correction factor 1.0273. With a beat frequency of 68.390 Hz we find the Sagnac rate as 70.26 Hz. This correction overshoots the expected value (68.95 Hz). Note that the Spectral lines do not fall off particularly linearly as the theory predicts; a result that has been noted on many data sets from CI both for single and combined beams.

The instantaneous frequency method gives indeterminate results for the data set taken from CII[1]. This is due to the first harmonic only being a few dB above the white noise floor. This makes the determination of the instantaneous phase inaccurate. Once the corrected Fourier spectrum is found for the instantaneous frequency the first harmonic is not visible and no ratio can be found between carrier and harmonics. The correction process cannot proceed.

E.2 The Direct Fourier Method

Here we determine a correction factor from the heights of the harmonics of the direct Fourier transform of the output data. The analysis gives the same theoret-

[1] The data set is stored in the file SS1B1 and comprises two channels of data which are interwoven. The two channels represent the output from a photomultiplier tube and a photodiode; the first 2^{14} samples of the photomultiplier output are used here the first sample having been taken at 29-Jun-1997 05:45:09

ical correction factor as for the instantaneous frequency method but the heights of the carrier and harmonics are found in a different way.

From Eq. (E.2) we know how the phase difference between the two counter propagating modes evolves with time. Our photomultiplier tube gives a signal linearly proportional to the intensity of the output light. With the counter propagating modes interfering we have

$$I \propto E^2 = E_{CW}^2 + E_{CCW}^2 + 2E_{CW}E_{CCW}\cos\Psi \tag{E.10}$$

and disregarding DC components and amplitudes we may write in general $I = \sin(\Psi(t)+\zeta)$. We choose $\zeta = \pi/2$ arbitrarily without altering the spectral content of the signal. A few manipulations give us

$$I = \frac{f_b \sin(2\alpha)}{f_s - f_l\cos(2\alpha)} \tag{E.11}$$

where $\alpha = \pi f_b t$. This is a purely odd function of time and hence the Fourier components are given by

$$F_n = \frac{2}{\tau}\int_0^\tau \sin(2\pi nt/\tau)\frac{f_b\sin(2\pi t/\tau)}{f_s - f_l\cos(2\pi t/\tau)}dt\ . \tag{E.12}$$

The evaluation of this integral follows in the same vein as for the instantaneous frequency method with the result

$$F_n = \frac{2f_b}{f_l}\left(\frac{f_s - f_b}{f_l}\right)^n \tag{E.13}$$

so that in the time domain the harmonics are expected to be in geometric progression with the same ratio as before $r = (f_s - f_b)/f_l$. This gives the same correction factor as in the instantaneous frequency analysis as shown in Eq. (E.9).

E.2.1 Results from Direct FFT Method

From the data set c96ap01c.sc1 three clearly visible harmonics for the direct Fourier transform of the time domain signal are seen. These first three harmonics are found 15.76 dB, 29.90 dB and 42.88 dB down from the carrier. Using the value for the first harmonic in the correction scheme outlined above gives a corrected output frequency of 72.12 Hz. This correction overshoots the theoretical value by a larger amount than the instantaneous frequency method.

The data taken from CII is also analysed and gives a first harmonic 44.19 dB down from the carrier. This harmonic is thus less than 1% of the carrier and suggests that the signal has little distortion and that the correction factor ought to be small. The factor turns out to be 1.000076 and corrects the output frequency from 79.622 Hz to 79.628 Hz. This therefore moves the output frequency away from the required frequency of 79.398 Hz.

Bibliography

[1] Robert Adler, *A study of Locking Phenomena in Oscillators*, *Proc. IEEE*, **61,** 1380–1385, 1973.

[2] H Akaike, *Fitting Autoregressive Models for Prediction*, *Ann. Inst. Statist. Math.*, **21,** 243–247, 1969.

[3] H Akaike, *Power Spectrum Estimation through Autoregression Model Fitting*, *Ann. Inst. Statist. Math.*, **21,** 407–419, 1969.

[4] David W Allan, *Statistics of Atomic Frequency Standards*, *Proc. IEEE*, **54,** 221–230, 1966.

[5] R Anderson, H R Bilger and G E Stedman, *"Sagnac" Effect: A Century of Earth-Rotated Interferometers*, *Am. J. Phys.*, **62,** 975–985, 1994.

[6] F Aronowitz, *Theory of a Travelling-Wave Optical Maser*, *Phys. Rev.*, **139,** A635–A646, 1965.

[7] F Aronowitz and R J Collins, *Lock-In and Intensity-Phase Interaction in the Ring Laser*, *J. Appl. Phys.*, **41,** 130–141, 1970.

[8] Fred Aronowitz and Wah L Lim, *PositIve Scale Factor Correction in the Laser Gyro*, *IEEE J. Quant. Electr.*, **QE-13,** 338–343, 1977.

[9] Frederick Aronowitz, *Laser Applications: The Laser Gyro*, Academic Press, New York, 1971.

[10] Frederick Aronowitz, *Single-Isotope Laser Gyro*, *Appl. Opt.*, **11,** 405–412, 1972.

[11] Frederick Aronowitz and R J Collins, *Mode Coupling Due to Backscattering in a He-Ne Travelling-Wave Ring Laser*, *Appl. Phys. Lett.*, **9,** 55–58, 1966.

[12] R H Barker, *Group Synchronization of Binary Digital Systems*, Butterworths, London, 1953, In *Communication Theory*: papers read at a symposium on "Applications of Communication Theory" held at the Institution of Electrical Engineers, London, September 22^{nd}-26^{th}, 1952, edited by Willis Jackson.

[13] J A Barnes, *Atomic Timekeeping and the Statistics of Precision Signal Generators*, *Proc. IEEE*, **54,** 207–220, 1966.

[14] S M Barnett and D T Pegg, *Phase in Quantum Optics*, *J. Phys. A*, **A 19,** 3849–3862, 1986.

[15] Markus Båth, *Mathematical Aspects of Seismology*, Elsevier, Amsterdam, 1968.

[16] Markus Båth, *Spectral Analysis in Geophysics*, Elsevier Scientific, Amsterdam, 1974.

[17] Ralph A Bergh, H C Lefevre and Herbert J Shaw, *An Overview of Fiber-Optic Gyroscopes*, *J. Lightwave Tech.*, **LT-2,** 91–107, 1984.

[18] H R Bilger and M R Sayeh, *Noise Phenomena in Ring Lasers*, presented at the 7th International Colloquium on Noise in Physical Systems, Montpelier, France, May 1983, printed in *Noise in Physical systems and 1/f noise* (Eds. M Savelli, G Lecoy and J P Nougier), Elsevier Science Publishers BV, 1983, p.325-328.

[19] H R Bilger, G E Stedman, Z Li, U Schreiber and M Schneider, *Ring Lasers for Geodesy*, *IEEE Trans. on Instr. and Meas.*, **44,** 468–470, 1995.

[20] H R Bilger, P V Wells and G E Stedman, *Origins of Fundamental Limits for Reflection Losses at Multilayer Dielectric Mirrors*, *Appl. Opt.*, **33,** 7390–7396, 1994.

[21] Hans R Bilger, G E Stedman, M P Poulton, C H Rowe, Ziyuan Li and P V Wells, *Ring laser for precision measurement of nonreciprocal phenomena*, *IEEE Trans. Instr. and Meas.*, **42,** 407–411, 1993.

[22] Max Born and Emil Wolf, *Principles of Optics*, Pergamon, Oxford, 5th edition, 1975, Optics.

[23] Ronald N Bracewell, *The Hartley Transform*, Clarendon Press, Oxford, 1986.

[24] Ronald N Bracewell, *The Fourier Transform*, *Sci. Am.*, **260,** 62–69, 1989.

[25] J P Burg, *Maximum Entropy Spectral Analysis*, in *Proceedings of the 37th Meeting of the Society of Exploration Geophysicists*, 1967, found in [27].

[26] S Chandrasekhar, *Stochastic Problems in Physics and Astronomy*, *Rev. Mod. Phys.*, **15,** 1–89, 1943, Appears reprinted in [123].

[27] Donald G Childers (Ed.), *Modern Spectrum Analysis*, IEEE Press, New York, 1978.

[28] W W Chow, J Gea-Banacloche, L M Pedrotti, V E Sanders, W Schleich and M O Scully, *The ring laser gyro*, *Rev. Mod. Phys.*, **57,** 61–104, 1985.

[29] Charles K Chui, *An introduction to wavelets*, Academic Press, Boston, 1992.

[30] F R Connor, *Noise*, Edward Arnold, London, 2nd edition, 1982.

[31] J W Cooley and J W Tukey, *An Algorithm for the Machine Calculation of Fourier Series*, *Mathematics of Computation*, **19,** 297–301, 1965.

[32] L Cooper and G E Stedman, *Axion detection by ring lasers*, *Phys. Lett. B*, **357,** 464–468, 1995.

[33] J D Cresser, *Quantum Noise in Ring Laser Gyros. III Approximate Analytic Results in Unlocked Region.*, *Phys. Rev. A*, **26,** 398–409, 1982.

[34] J D Cresser, *Theory of the Spectrum of the Quantised Light Field*, *Phys. Rep.*, **94,** 47–110, 1983.

[35] J D Cresser, D Hammonds, W H Louisell, P Meystre and H Risken, *Quantum Noise in Ring Laser Gyros. II Numerical Results*, *Phys. Rev. A*, **25,** 2226–2234, 1982.

[36] J D Cresser, W H Louisell, P Meystre, W Schleich and M O Scully, *Quantum Noise in Ring Laser Gyros. I Theoretical Formulation of the Problem.*, *Phys. Rev. A*, **25,** 2214–2225, 1982.

[37] Germund Dahlquist and Åke Björck, *Numerical Methods*, Prentice-Hall, Englewood Cliffs, New Jersey, 1974, Translated by Ned Anderson.

[38] Giacomo M D'ariano, Chiara Macchiavello and Matteo G. A. Paris, *Feasible phase detection with ideal sensitivity*, *J. Phys. A*, **29,** 5605–5610, 1996.

[39] Giacomo M D'ariano and Matteo G. A. Paris, *Arbitrary precision in multipath interferometry*, *Phys. Rev. A*, **55,** 2267–2271, 1997.

[40] P A M Dirac, *The Quantum Theory of the Emission and Absorption of Radiation*, *Proc. Roy. Soc.*, **A 114,** 243–265, 1927.

[41] Terry A Dorschner, Hermann A Haus, Michael Holz, Irl W Smith and Hermann Statz, *Laser Gyro at Quantum Limit*, *IEEE J. Quant. Electr.*, **QE-16,** 1376–1379, 1980.

[42] Robert W Dunn, *Design of a triangular active ring laser 13 m on a side*, *Appl. Opt.*, **37,** 6405–6409, 1998.

[43] C D Ellyett and G J Fraser, *The Influence of Noise on Radar Meteor Observations*, *Austr. J. Phys.*, **8,** 273–278, 1955.

[44] Paul M Embree and Bruce Kimble, *C Language algorithms for digital signal processing*, Prentice-Hall, Englewood Cliffs, New Jersey, 1991.

[45] C Etrich, Paul Mandel, R Centeno Neelen, R J C Spreeuw and J P Woerdman, *Dynamics of a Ring-Laser Gyroscope with Backscattering*, *Phys. Rev. A*, **46,** 525–536, 1992.

[46] M Freyberger and A M Herkommer, *Probing a Quantum State via Atomic Deflection*, *Phys. Rev. Lett.*, **72,** 1952–1955, 1994.

[47] D Gabor, *Theory of Communication*, *J. Inst. Electr. Engrs.*, **93,** 429, 1946.

[48] John C Garrison and Jack Wong, *Canonically Conjugate Pairs, Uncertainty Relations, and Phase Operators*, *J. Math. Phys.*, **11,** 2242–2249, 1970.

[49] H Gerhardt, U Büchler and G Litfin, *Phase Measurement of a Microscopic Radiation Field*, *Phys. Lett.*, **49A,** 119–120, 1974.

[50] Christopher C Gerry and Kenneth E Urbanski, *Hermitian Phase-Difference Operator Analysis of Microscopic Radiation-Field Measurements*, *Phys. Rev. A*, **42,** 662–664, 1990.

[51] Roy J Glauber, *Coherent and Incoherent States of the Radiation Field*, *Phys. Rev.*, **131,** 2766–2788, 1963.

[52] Roy J Glauber, *The Quantum Theory of Optical Coherence*, *Phys. Rev.*, **130,** 2529–2539, 1963.

[53] Philip Goldberg, Peter W Milonni and Bala Sundaram, *Theory of the fundamental laser linewidth*, *Phys. Rev. A*, **44,** 1969–1985, 1991.

[54] R W Hamming, *Digital Filters*, Prentice-Hall, Englewood Cliffs, New Jersey, 1977.

[55] N Harris, *Modern Vacuum Practice*, McGraw-Hill, London, 1989.

[56] R. V. L. Hartley, *A More Symmetrical Fourier Analysis Applied to Transmission Problems*, *Proc. Inst. Radio Eng.*, **30,** 144–150, 1942.

[57] Hermann A Haus, Hermann Statz and Irl W Smith, *Frequency Locking of Modes in a Ring Laser*, *IEEE J. Quant. Elec.*, **QE-21,** 78–85, 1985.

[58] Richard W. Henry and Sharon C. Glotzer, *A squeezed-state primer*, *Am. J. Phys.*, **56,** 318–328, 1988.

[59] Paul Horowitz and Winfield Hill, *The art of electronics*, Cambridge University Press, 2nd edition, 1989.

[60] T J Hutchings, J Winocur and W L Zingery, 5th *Symp. Unconventional Inertial Sensors*, Naval Applied Science Laboratory, Brooklyn, New York, 1969.

[61] C Huygens, *J. des Scavans*, no. 11 Mar. 16; no. 12, Mar. 23, 1665, Huygens' notebook is reprinted in Oeuvres Complètes de Christiaan Huygens, Societé Hollandaise des Sciences, vol. 17, p. 185, 1888–1950.

[62] E T Jaynes, *Prior Probabilities*, *IEEE Trans. Systems Sci. Cybern.*, **SEC-4,** 227–241, 1968.

[63] Garside B K, *Mode Spectra in Ring and Normal Lasers*, *IEEE J. Quant. Electr.*, **QE-4,** 940–948, 1968.

[64] Steven M Kay and Stanley Lawrence Marple, *Spectrum Analysis - A Modern Perspective*, *Proc. IEEE*, **69,** 1380–1419, 1981.

[65] Stanislav B Kesler, *Modern Spectrum Analysis II*, IEEE Press, New York, 1986.

[66] B. Tom King, *An Einstein Addition Law for Nonparallel Boosts Using the Geometric Algebra of Space-Time*, *Found. Phys.*, **25,** 1741–1755, 1995.

[67] B Tom King, *Reflection Optics, an Einstein Addition Law for Nonparallel Boosts, and the Geometric Algebra of Space-Time*, *J. Opt. Soc. Am. A*, **12,** 773–779, 1995.

[68] W. Christopher Lang and Kyle Forinash, *Time-frequency analysis with the continuous wavelet transform*, *Am. J. Phys.*, **66,** 794–797, 1998.

[69] American Radio Relay League, *The Radio Amateur's Handbook*, American Radio Relay League, Newington, Conn., 54th edition, 1977.

[70] H C Lefevre, Y Bourbin, H Graindorge and H J Arditty, *Review of fiber optic gyroscopes*, in *The Max Born Centenary Conference*, **SPIE-369**, 385–394, Washington, 1982, Soc. Photo-Optical Instr. Eng.

[71] Jean-Marc Lévy-Leblond, *Who Is Afraid of Nonhermitian Operators? A Quantum Description of Angle and Phase*, *Ann. Phys.*, **101,** 319–341, 1976.

[72] Rodney Loudon, *The Quantum Theory of Light*, Clarendon Press, Oxford, 2nd edition, 1983.

[73] Robert Lynch, *Fluctuations of the Barnett-Pegg Phase Operator in a Coherent State*, *Phys. Rev. A*, **41,** 2841–2843, 1990.

[74] Robert Lynch, *The Quantum Phase Problem: A Critical Review*, *Phys. Rep.*, **256,** 367–436, 1995.

[75] S Lawrence Marple, *Digital Spectrum Analysis: with Applications*, Prentice Hall, Englewood Cliffs, New Jersey, 1987.

[76] J V Martinez, *Influence of Plasma-Tube-Surface Interaction on the Helium-Neon Laser Lifetime*, *J. Appl. Phys.*, **37,** 4477–4483, 1966.

[77] Duncan P McLeod, *Working Title: Transverse Mode Studies in He-Ne Lasers and Seismic Effects in Ring Lasers.*, PhD thesis, University of Canterbury, Christchurch, New Zealand, 2000.

[78] Prabhakar S Naidu, *Modern Spectrum Analysis of Time Series*, CRC Press, Boca Raton, 1996.

[79] Lawrence C Ng and Darryll J Pines, *Characterisation of Ring Laser Gyro Performance Using the Allan Variance Method*, *J. Guidance*, **20,** 211–214, 1996.

[80] Michael Martin Nieto, *Phase-Difference Operator Analysis of Microscopic Radiation Field Measurements*, *Phys. Lett.*, **60A,** 401–403, 1977.

[81] J W Noh, A Fougères and L Mandel, *Measurement of the Quantum Phase by Photon Counting*, *Phys. Rev. Lett.*, **67,** 1426–1429, 1991.

[82] J W Noh, A Fougères and L Mandel, *Further Investigations of the Operationally Defined Quantum Phase*, *Phys. Rev. A*, **46,** 2840–2852, 1992.

[83] J W Noh, A Fougères and L Mandel, *Operational Approach to the Phase of a Quantum Field*, *Phys. Rev A*, **45,** 424–442, 1992.

[84] Tomáš Opatrný, *Number-Phase uncertainty relations*, *J. Phys. A*, **28,** 6961–6975, 1995.

[85] Alan V Oppenheim and Ronald W Schafer, *Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, New Jersey, 1975.

[86] Philip F Panter, *Modulation, Noise and Spectral Analysis*, McGraw-Hill, New York, 1965.

[87] H Paul, *Phase of Microscopic Electromagnetic Field and its Measurement*, *Fortschr. Phys.*, **22,** 657, 1974.

[88] D T Pegg and S M Barnett, *Unitary Phase Operator in Quantum Mechanics*, *Europhys. Lett.*, **6,** 483–487, 1988.

[89] David T. Pegg and Stephen M. Barnett, *Tutorial Review: Quantum Optical Phase*, *J. Mod. Opt.*, **44,** 225–264, 1997.

[90] L Pesquera and R Blanco, *Analysis of the Intensity Correlation Functions of gas Ring Lasers with Backscattering*, *Opt. Comm.*, **74,** 102–106, 1989.

[91] R L Pfleegor and L Mandel, *Interference of Independent Photon Beams*, *Phys. Rev.*, **159,** 1084–1088, 1967.

[92] R L Pfleegor and L Mandel, *Further Experiments on Interference of Independent Photon Beams at Low Light Levels*, *J. Opt. Soc. Am.*, **58,** 946–950, 1968.

[93] V N Popov and V S Yarunin, *Quantum and Quasi-Classical States of the Photon Phase Operator*, *J. Mod. Opt.*, **39,** 1525–1531, 1992.

[94] W H Press, B P Flannery, S A Teukolsky and W T Vettering, *Numerical Recipes in C*, Cambridge University Press, 1988.

[95] W V Prestwich, T J Kennett and F W Kus, *The Statistical Properties of Allan Variance*, *Can. J. Phys.*, **69,** 1405–1415, 1991.

[96] M.B. Priestley, *Spectral Analysis and Time Series*, Academic Press, London, 1981.

[97] L R Rabiner and R W Schafer, *On the Behaviour of Minimax FIR Digital Hilbert Transformers*, *Bell Syst. Tech. J.*, **53,** 363–390, 1974.

[98] H Risken, *The Fokker-Planck Equation: Methods of Solution and Applications*, Springer-Verlag, Berlin, Germany, 2nd edition, 1989.

[99] C H Rowe, U Schreiber, S J Cooper, B Tom King, M Poulton and G E Stedman, *Design and operation of a very large ring laser gyroscope*, *Appl. Opt.*, **38,** April, 1999.

[100] Ch. Salomon, D Hils and J L Hall, *Laser stabilisation at the millihertz level*, *J. Opt. Soc. Am. B*, **5,** 1576–1587, 1988.

[101] Murray Sargent, Marlan O Scully and Willis E Lamb, *Laser Physics*, Addison-Wesley, Reading, Mass., 1974.

[102] W P Schleich, D F Walls and J A Wheeler, *Area of Overlap and Interference in Phase Space Versus Wigner Pseudoprobabilities*, *Phys. Rev. A*, **38,** 1177–1186, 1988.

[103] Claude E Shannon and Warren Weaver, *The Mathematical Theory of Communication*, University of Illinois Press, Urbana, 1971.

[104] Anthony E Siegman, *Lasers*, University Science Books, Mill Valley, California, 1986.

[105] P W Smith, *Linewidth and Saturation Parameters for the* 6328*A Transition in He-Ne Laser*, *J. Appl. Phys*, **37,** 2089–2093, 1966.

[106] D E Smylie, G K C Clarke and T J Ulrych, *Analysis of irregularities in the Earth's rotation*, Academic Press, 1973, in Berni Adler, Sidney Fernbach and Manuel Rotenberg (Eds.) *Methods in computational Physics* vol. 13, pp. 391-430.

[107] Hermann Statz, Terry A Dorschner, Michael Holtz and Irl Smith, *Laser Handbook vol.4 : The Multioscillator Ring Laser Gyroscope*, Elsevier, North-Holland, Amsterdam, 1985.

[108] G E Stedman, *Ring-laser tests of fundamental physics and geophysics*, *Rep. Prog. Phys.*, **60**, 615–688, 1997.

[109] G E Stedman, H R Bilger, Li Ziyuan, M P Poulton, C H Rowe, I Vetharaniam and P V Wells, *Canterbury Ring Laser and Tests for Nonreciprocal Phenomena*, *Austr. J. of Phys.*, **46**, 87–101, 1993.

[110] G E Stedman, M T Johnsson, Z Li, C H Rowe and H R Bilger, *T violation and microhertz resolution in a ring laser*, *Opt. Lett.*, **20**, 324–326, 1995.

[111] G E Stedman, Z Li and H R Bilger, *Sideband analysis and seismic detection in a large ring laser*, *Appl. Opt.*, **34**, 5375–5385, 1995.

[112] G E Stedman, Z Li, C H Rowe and A D McGregor, *Harmonic analysis in a large ring laser with backscatter-induced pulling*, *Phys. Rev. A*, **51**, 4944–4958, 1995.

[113] L Susskind and J Glogower, *Quantum Mechanical Phase and Time Operator*, *Physics*, **1**, 49, 1964.

[114] Harry L Van Trees, *Detection, Estimation and Modulation Theory*, John Wiley, New York, 1971.

[115] Y K Tsui and Michael F Reid, *Unitary and Hermitian Phase Operators for the Electromagnetic Field*, *Phys. Rev. A*, **46**, 549–554, 1992.

[116] Yee Kin Tsui, *Quantum Phase Opertaors, Theory and Applications*, Master's thesis, University of Hong Kong, 1992.

[117] Tad J Ulrych and Thomas N Bishop, *Maximum Entropy Spectral Analysis and Autoregressive Decomposition*, *Rev. Geophysics and Space Phys.*, **13**, 183–200, 1975.

[118] H C van de Hulst, *Light scattering by small particles*, Wiley, New York, 1957.

[119] A VanDenBos, *Alternative Interpretation of Maximum Entropy Spectral Analysis*, *IEEE Trans. Inform. Theory*, **IT-17**, 493–494, 1971, Found in [27].

[120] Joseph T Verdeyen, *Laser Electronics*, Prentice-Hall, New Jersey, second edition edition, 1989.

[121] W Vogel and W Schleich, *Phase Distribution of a Quantum State Without Using Phase States*, *Phys. Rev. A*, **44**, 7642–7646, 1991.

[122] J G Wade, *Signal Coding and Processing*, Cambridge Universty Press, Cambridge, 1994.

[123] N Wax (Ed.), *Selected papers on noise and stochastic processes*, Dover, New York, 1954.

[124] Robert C. Weast (Ed.), *Handbook of Chemistry and Physics*, CRC Press, Boca Raton, Florida, 64^{th} edition, 1984.

[125] Peter D Welch, *The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Averaging Over Short, Modified Periodograms*, *IEEE Trans. Audio and Electroacoust.*, **AU-15,,** 70–73, 1967.

[126] J R Wilkinson, *Ring Lasers*, *Prog. Quant. Electr.*, **11,** 1–103, 1987.

[127] Kawakatsu Yamada, Tetsuo Yamazaki, Takashi Shimizu, Norihiro Sei and Tomohisa Mikado, *Plasma Treatment for Restoration of Dielectric Multilayer Mirrors in Short-Wavelength Free-Electron Lasers*, *Appl. Opt.*, **34,** 4261–4265, 1995.

[128] Li Ziyuan, *Optical Supercavity and Precision Ring Laser Measurements*, PhD thesis, University of Canterbury, Christchurch, New Zealand, 1993.

[129] Li Ziyuan, R G T Bennett and G E Stedman, *Swept-frequency induced optical cavity ringing*, *Opt. Comm.*, **86,** 51–56, 1991.

[130] Li Ziyuan, G E Stedman and H R Bilger, *Asymmetric response profile of a scanning Fabry-Perot interferometer*, *Opt. Comm.*, **100,** 240–247, 1993.